A COMPARATIVE STUDY
of
SOME FACTOR-ANALYTIC TECHNIQUES

by

A.E. Maxwell
M.A., (Dublin), B.Sc., (Lond.), B.Ed., (Edin.)

Thesis presented in fulfilment of the requirements for the Degree of Doctor of Philosophy, Edinburgh University.

April, 1957.

ACKNOWLEDGEMENTS

Sincerest thanks are due to my supervisors, Professor J. Drever, Dr. B. Semeonoff and Dr. D.N. Lawley. To Professor Drever who, knowing that my thesis was going to deal largely with statistical problems, to my advantage never allowed me to forget the psychological aspects of the subject. To Dr. Semeonoff, who patiently bore with me especially through the initial stages of the work and helped me to achieve some unity in the great diversity of controversial topics which I proposed to discuss. And finally to Dr. Lawley, of the Mathematical Institute, Edinburgh University, who acted as statistical consultant and on whom the main burden of directing my work fell; I will always be indebted to him for his very expert guidance.

My sincere thanks are also due to Professor H.J. Eysenck of the Institute of Psychiatry, London University, for permission to use - as illustrative material - data which passed through my hands in the statistical laboratory of the Institute. I am indebted too to Mr. W.G. Emmett and others whose published data I have quoted at some length.

Finally I owe a great deal to my wife, who typed the thesis, for her diligence in eliminating grammatical errors from the script and for drawing attention to obscurities in thought and carelessness in style. Were it not for her this thesis would be even more difficult to read.

C O N T E N T S

TABLES

TABLES (continued)

Chapter I

INTRODUCTION

It is no secret that many people have misgivings about factor analysis. They query its claims, doubt the statistical validity of many of its basic techniques, and are unconvinced of its practical value. The writer, while himself actively engaged in factor-analytic work, has at times had all of these misgivings and this study was undertaken in the hope of clarifying his own thoughts on the subject. The stress throughout the study, except in chapter II which refers briefly to the growth of factor analysis and deals with basic logical and psychological issues, is on the statistical and methodological aspects of the subject. If we agree with Eysenck (1953, p.105) that these aspects are much less fundamental and much more easily settled than questions about the logic of factor analysis then the writer has chosen the less difficult task, yet it will be shown that even here, with some notable exceptions (Lawley, 1940; 1949), the difference between nescience and knowledge is not always great.

However, it is possible to be over-critical even of one's own subject and although many of the basic problems of factor analytic <u>theory</u> were inadequately dealt with in the factorial procedures employed, plain common sense on the part of research workers seems to have compensated for the handicap with the result that many valuable studies have been reported. Still the factor analyst using the more common and approximate techniques finds himself in a weak position when faced by an adversary, for while his arguments regarding the reasonableness of the factor concept may be plausible, he is defenceless when asked how he tests a matrix of 'residuals' for significance, or decides when a 'loading' is sufficiently small to be considered to be zero. On the practical side it may be shown that such shortcomings are often of no great import and the property of being significant does not endow a factor with meaning. But, while defects in the popularly used factor techniques remain, the factor analyst's stock will continue to be low in the eyes of others, and his work will be viewed with suspicion by his colleagues.

The Literature.

Factor analysis is still young but already the literature on it is very extensive. No attempt will be made here to survey it since many comprehensive surveys and authorative works on the subject by both British and American writers already exist. Passing reference, however, to the leading British writers will help to get the subject into perspective. Thomson (1939) in his book "The factorial analysis of human ability", gives a very clear exposition and able assessment of the more commonly used factorial techniques. Burt's "Factors of the mind", (1940) is especially valuable for its closely reasoned treatment of the psychological, logical, and metaphysical aspects of the subject. In "The structure of human abilities", Vernon (1950) evaluates with masterly ease and dexterity the contribution of factor analysis to our understanding of cognitive ability.

The use of factorial techniques too, as all the writers mentioned agree, is not confined to research in the cognitive field, and in this country BURT, (1947-48) and Eysenck (1947; 1952) in particular, has demonstrated how factor analysis can be used to advantage in personality studies. Indeed ~~he~~ EYSENCK claims that "in the present stage of development of psychology, factor analysis is an indispensable method of taxonomic and nosological research". If such a claim is justified every effort should be made to eliminate defects from methods in popular use and to encourage the use of efficient procedures - like Lawley's maximum likelihood method - where such exist.

Electronic computing aids.

In the past there has been a tendency to neglect more rigorous factorial procedures on the grounds that the laborious calculations involved were not justified. This excuse may be no longer valid. Three years ago when this study was commenced a few factor analysts in America, for example Anderson at Chicago (1955), and Wrigley and Neuhaus at Illinois (1954), had already had access to electronic computing facilities. But such facilities are not confined to the States; they also exist in Britain if a search of one's locality is made. The psychologist still tends to be awed by the mention of electronic computers, but he may rest assured that to the electronic engineer his calculations will appear straightforward. The main handicap is generally not one of 'programming' but of the storage capacity of the computing machine which for factorial work is generally required to be big.

The promise of relief from heavy calculations on desk machines has influenced the recommendations put forward in this study. The author has not hesitated to recommend tests and techniques involving laborious calculations where their use would help to clarify the results of an investigation. A good example is the test for the significance of the difference between two variance-covariance matrices given in chapter VI. This involves evaluating the determinants of three large matrices - a task which normally would be quite out of the question.

An old controversy.

The advent of the electronic computer, however, has re-opened an old controversy, since it enables the experimenter to resolve a correlation or covariance matrix into its principal components with considerable ease. This has led some psychologists (Wrigley and Neuhaus, 1955; Warburton, 1954) to employ this type of analysis in preference to the more generally accepted factor methods, lured, as they admit, by a certain arithmetic tidyness which the principal component method possesses. As a result it now becomes desirable to re-evaluate the latter method, a task which is rendered difficult because it is the method recommended by some of our leading mathematical statisticians (Kendall, 1950, p.63)* who seem not to be aware of possible psychological objections. As a result the psychologist if often placed in a dilemma, for while he is in need of the statistician's help he fears lest, by accepting it, he should be led to betray his psychological insight.

---

* Since this paragraph was written a joint note entitled, "The Principles of Factor Analysis", by Prof. M.G. Kendall and Dr. D.N. Lawley has appeared (J.Roy. Stat. Soc., B-series, 1956) which should go a long way towards dispelling the confusion which exists regarding factor analysis and principal component analysis. In this note the fundamental difference between the two models is clearly stated. As a result it is likely that the factor model will in future to accepted - and rightly so - as a distinct model in its own right.

The aim of the study.

It appears then that if this study is to have any value it must aim at a rapprochement between psychologist and statistician. The psychological concepts which have inspired the development of factor analysis must be critically reviewed. Then the types of model most suited to represent these concepts at the concrete practical level must be chosen. The sampling theory necessary for a valid use of such models must then be stressed, and a statistical assessment of models in present use made.

With the aim of fulfilling this programme a comparative approach has been chosen. But the number and variety of existing factorial techniques is so great that at the outset a choice of those which it was thought would lead to the most useful comparison and most pertinent discussion was necessary. Accordingly, the Centroid, or Simple Summation method, was chosen as being that most widely used. The Principal Component, or Principal Axes method was chosen as representing 'the statistician's choice'. Lastly, the Maximum Likelihood method was chosen as perhaps the only one so far to appear which meets the psychologist's acclaimed needs as well as the sterner dictates of sound statistical theory.

All these are methods of extracting factors. They are considered in chapter III, where several correlation matrices have been factorized by two or more methods and the results presented for comparison. As well as the problem of factor extraction there is the equally important problem of deciding how many factors to extract. In the past this problem has been responsible for a great deal of controversy. Chapter IV is devoted to a discussion of it, and to an examination of some of the empirical rules which have been put forward for deciding when to stop factoring. Results given by these empirical rules are compared with those obtained by the use of efficient tests.

To reiterate, comparisons of this kind are a feature of this study, for while it may be argued that in future with improved computing facilities research workers are likely to employ only efficient techniques in their basic research, it will be valuable to know just how efficient the older methods are. This will help when making comparisons with earlier studies and for making quick assessments of the possible value of analysing virgin data.

In chapter V the vexed question of the interpretation of factors is considered. Here a number of the more recent suggestions for obtaining unique or objective solutions are employed, and the results they give are evaluated against the more orthodox precepts of common sense and experience. At this stage a plea is made for the more careful design of factorial experiments.

In chapter VI the problem of factorial invariance and of the comparison of factors obtained from replicated experiments is dealt with. Empirical methods for comparing factors suggested by Barlow and Burt, and by Ahmavaara are reviewed and evaluated and an approach to the problem along more orthodox statistical lines is outlined and illustrated. The latter approach, however, does not prove fruitful from a practical viewpoint and an alternative approach - made possible by recent work by Howe and by Lawley - is recommended.

Finally, in chapter VII, the main conclusions reached in the thesis are summarised and the present status of factor analysis as a branch of statistical theory assessed. In conclusion it is recommended that the psychologist who has doubts about the value of factorial techniques should do a factorial study of his own. Only in this way will he have an opportunity of facing for himself the problems referred to, and of passing from theory to the immanence of a real experience.

Chapter II

FACTOR ANALYSIS AS INTERACTION BETWEEN MATHEMATICAL AND PSYCHOLOGICAL IDEAS

It is easy to become so engrossed in the mathematical and statistical aspects of factor analysis that one loses sight of the psychological aspects, and since this thesis is concerned in particular with statistical issues it may be well at the outset to outline briefly the psychological problems on which factor analysis tries to throw light.

The aims of factor analysis

Factor analysis starts with a matrix of covariances or correlations between a set of variables. In psychological work these variables generally consist of cognitive or attainment tests, attitude questionnaires, psycho-physical readings or other measures of human behaviour. The chief aim of the factor analyst is to test some hypothesis concerning his data (e.g., that there is only a single cognitive factor), but often - if the study is of an exploratory nature - he may find himself simply examining his matrix to discover if any systematic patterns exist among the inter-correlations. In either case he will wish to give a concise statistical description of the matrix concerned and to verify hypotheses by the application of suitable significance tests. Proceeding in this way he hopes to get a better understanding of human behaviour and of the variables being employed to measure it.

Take as an example the correlation matrix in table I. The first five variables are measures of neurotic tendency, while

TABLE I

Correlation Matrix Showing Two Clusters

Tests	1	2	3	4	5	6	8	10
1		475	297	134	205	-133	-222	~~-212~~ -122
2			479	424	448	-517	-379	-399
3				453	381	-305	-334	-161
4					321	-151	-106	-190
5						-271	-293	-091
6							570	568
8								335
10								

the remaining three are measures of intelligence. The members of the first set form a cluster in that they correlate positively and highly with each other. So, too, do the members of the remaining set. But any variable from the first set correlates negatively, (and in general only to a small extent) with any variable from the second set. There appears then to be a contrast between these two sets of variables. The problem is how to talk objectively about this contrast and also about the similarity between the variables within each set.

Caveat on 'cause' and 'meaning'.

Now if we turn to table II the same eight variables appear in conjunction with five other variables. The situation is now more complex. For one thing the number of intercorrelations, being a quadratic function of the number of variables employed, has been greatly increased. A cursory examination of the new matrix shows that variables 11 and 13 appear to 'go with' 6, 8, 10, in that they correlate positively with them and negatively with numbers 1 to 5. On the other hand variables 12, 14, 15, for similar reasons, seem to 'go with' 1 to 5, although the indications are not so pronounced. So far, however, we have not been told what variables 11 to 15 are, and clearly we are not in a position to make any inference about their nature from the observations just made about how they correlate with those we know. Here we recall the well known fact that

TABLE II

Supplementing the Clusters in Table I with Additional Tests

Tests	1	2	3	4	5	6	8	10	11	12	13	14	15
1		475	297	134	205	-133	-222	-122	-097	376	-083	222	240
2			479	424	448	-517	-379	-399	-291	302	-209	125	249
3				453	381	-305	-334	-162	-284	300	-283	055	141
4					321	-151	-106	-190	-124	075	-252	173	152
5						-271	-294	-091	-127	049	-191	047	088
6							570	568	182	-172	130	034	-136
8								335	254	-352	185	-086	-249
10									244	-063	051	-087	-192
11										-176	208	-293	-234
12											-204	069	235
13												-123	069
14													694
15													

whereas a significant correlation between two variables tells us that there is an association between them it tells us nothing of the nature, 'cause' or 'meaning' of this association. It follows then that the task of the psychologist is not just one of examining and describing his matrix from a statistical viewpoint. Once a cluster of highly correlating variables has been isolated some reason for its existence must be sought. On this point the analysis of the data into 'factors' offers no help. 'Causal' explanations must be sustained on other than statistical grounds. Here a passage from Burt's book (1940, p.68) will repay contemplation. Discussing factors as providing 'causal explanations' he states - "to borrow Thomson's language I would say, not merely that the 'causal entities' discovered by factor analysis <u>may</u> be 'things we already know of in other connections', but that, if they are entities at all, they <u>must</u> be 'things we already know of in other connections,'" or at least things that we have antecedent reasons to postulate as probable or as convenient; their existence is in no way attested by the process of factorization". This statement does not, of course, preclude the possibility that factor analysis may be suggestive of new hypotheses - which, as Eysenck (1953) points out is one of its possible functions. It only asserts that the evidence on which such hypotheses are accepted as new must be <u>a priori</u> evidence and is not a consequence of the analysis.

The 'structure' of a correlation matrix.

In small matrices like the two just examined it is often possible to discover at sight any patterns which exist among the correlations, and in this way some idea of the 'structure' of the matrix is obtained. The task too is made much easier, (Burt, 1950, pp.45-6) if the experimenter, from his prior knowledge of the variables, groups like with like when writing out the matrix. But overlapping patterns between a set of variables may exist and in a large matrix may escape the eye. It is here that the special techniques which the factor analyst has devised prove indispensable. The methods available are reviewed by Thomson (1939, 5th. ed. p.20) and Burt (1950, p.40), but more will be said of this later.

The statistical problem of how best to describe a correlation matrix was one which naturally arose once the use of correlational techniques had become widespread. In an article entitled "Alternative methods of factor analysis and their relations to Pearson's method of 'principal axes'", Burt (1949) tells us of the early attacks on the problem by Galton, Pearson and others, and of Pearson's conclusion (1901) that what was required was the resolution of the matrix into uncorrelated components. This approach resulted in the principal axes method of Pearson and later in Hotelling's (1933) better known principal component method. In both methods the original correlated test vectors are transformed into an equal number of orthogonal components. From the historical viewpoint Burt and his disciples lay most stress on this approach to the problem, although there is another approach - that of Spearman - which has appealed more to other psychologists.

Spearman's contribution.

It had long been observed that correlations between tests of cognitive ability tended to be positive. What Spearman was the first to notice, as Thomson records (1929), was that the 'tetrad-differences' calculated from the entries in a matrix of correlations between such tests tended to be zero. An artificial example will help us to get the argument clear.

Suppose that the correlations between four cognitive tests are as follows:-

Tests.	1	2	3	4
1		.30	.24	.18
2	.30		.20	.15
3	.24	.20		.12
4	.18	.15	.12	
Total	.72	.65	.56	.45

Now if we take from this table any four entries which form a square or rectangle, for example the four entries in the south-west corner, we find that the value of the tetrad-difference

$$(.24 \times .15) - (.18 \times .20)$$

is zero. In this artificial example the value is exactly zero, but with coefficients obtained in practice Spearman noticed that there was a strong tendency for this to be the case.

The tests in our example have been arranged in descending order as regards the size of their correlations with the other tests. This is shown by the totals of the columns in the last row of the table. In this way the 'hierarchical order' of the tests, which again is a reflection of the tendency towards zero of the tetrad-differences, is clearly displayed. This 'hierarchical order', Spearman explained by the hypothesis that all the correlations were due to a single 'factor' common to all the tests, but present to the greatest extent in the test at the head of the hierarchy. For cognitive tests this factor was his famous 'g', and every psychologist is aware of the controversy to which Spearman's explanation gave rise, and of the speculations which arose as to the nature, and 'reality' of 'g'.

Spearman's 'common' factor was postulated to account for the covariation of the tests. To complete the picture it was necessary to postulate a further factor for each test which was 'specific' to it and which accounted for the remainder of the variance.

This analysis of the variance of each test into two parts suggests how the diagonal cells of an 'hierarchical matrix' may be filled. When describing a correlation matrix it is the co-variance or common variance of the tests in which we are interested. It is reasonable then to enter in the diagonal cells of the matrix only that part of the variance of each test which it shares with the others. In other words the diagonal cells should be filled in such a way as to preserve the hierarchical structure. Once this step is taken it is possible to calculate what the correlations of each test with the hypothetical 'factor' are. These correlations, called 'loadings' in our artificial example are, .6, .5, .4, and .3. Their special merit is that they can be employed to reproduce the original correlations and so are in keeping with Spearman's contention that the correlations are due to one 'common factor'.

Further developments of the early ideas.

These then are the two roots from which factor analysis sprang. Pearson's approach was further developed by Hotelling (1933), who supplied an iterative method for calculating the latent roots and vectors of the matrix. Spearman's "two factor theory", as it was called, was soon found to be an oversimplification of the facts, and it was replaced by a multiple-factor theory. In this development the work of Thurstone (1947) is of special interest as it is in one major respect a generalization of Spearman's ideas. Multiple factor methods were also developed by Burt (1915) (1940) following Pearson's approach, but using 'reduced self-correlations' in the diagonal cells of the matrix and thereby reducing the number of components obtained.

The fact that the tetrad-differences are zero in the side matrix of an hierarchical example is equivalent in mathematical language to saying that the rank of the side matrix is one, and we have seen that Spearman postulated one factor to account for the co-variation between the variables. But if the rank of the side matrix were two it seemed natural to generalise Spearman's idea and postulate that two 'factors' would be required to account for the co-variation between the tests, and so on. This briefly was Thurstone's contribution, as far as the extraction of factors was concerned.

The centroid and principal component methods.

Now that we have reviewed very briefly early developments it will be well to look in more detail at the principal component and centroid methods.

The first thing to notice about the principal axes or principal components method is - as Bartlett (1953) has recently stressed - that it is purely an empirical method for 'breaking down' a correlation (or covariance) matrix. When employed with unities in its diagonal cells, as is customary, a set of orthogonal components or axes equal in number to the number of variables employed can be obtained. These components may then be used to reconstruct either the correlation matrix or the original test scores and if all the components are used these can be reconstructed exactly. The procedure can be thought of in terms of a geometrical model. If there are n subjects and p tests then the scores of the subjects can be represented by n points in p dimensions and the full analysis results in an orthogonal transformation of the p axes to new positions where the first axis evolving from the calculations is the major axis of the multi-dimensional ellipsoid of test points, the second axis is the second major axis of the ellipsoid, and so on.

The psychologist must note that no search for 'patterns' within the correlation matrix is made and no psychological hypotheses are involved. The components that result from such an analysis have no significance beyond themselves.

When we turn to the Centroid Method of extracting factors we find that though it is, in the sense stated earlier, a generalization of Spearman's approach, it bears perhaps more resemblance to the principal component method. The model used is again multi-dimensional. The variables are represented by vectors intersecting at their means, the cosine of the angle between each pair being equal to the correlation between them. Again too there is no direct search for pattern within the correlation matrix and the factors arrived at - unlike the Spearman common factor - are not unique.

Details of how the centroid factors are obtained will not be given here as the procedure is a bit unwieldy. It is, however, well described in Thurstone's book (1947) and again by Thomson (1939). Some of the most serious difficulties encountered when using it are dealt with in the practical examples given in the next chapter. It is worth mentioning here, however, that while the method can be used with unities in the diagonal cells of the correlation matrix the ~~procedure~~ modification is unusual. Customarily one fills the diagonals by values less than unity which conform approximately to the rank of the side matrix and which have to be found by iterative processes. This procedure is directly aimed at reproducing the correlation co-efficients by as few factors as possible. The advisability of doing so has been seriously questioned - in particular by Thomson (1939a), but we will return later to this problem.

The biggest contrast between the two methods just discussed can now be seen. At one extreme we have the principal component method giving a maximum number of 'common factors', that is as many as there are variables. The centroid method with reduced diagonal entries, or 'communalities', conforming to the rank of the side matrix, gives a minimum number. The ideal solution it is sometimes claimed (Burt and Banks, 1954) may lie in a compromise between the two extremes and many workers feel that Burt's 'group factor' approach is the most satisfactory. The latter is, at present, worked out only in an approximate way and would require more rigorous treatment to be fully acceptable from a statistical viewpoint.

Factors as 'causal entities'

'Causal explanations' of factors have already been touched on with reference to clusters in a correlation matrix. The need to mention the topic again arises in connection with the rank of a matrix, and affords an opportunity to underline the warnings given earlier.

When the rank of the side matrix is low and the covariation between a set of variables can be accounted for by a relatively small number of factors - ~~like~~ MUCH AS the hierarchical matrix can be accounted for by one - there is a tendency to think that this mathematical fact in some way endows the factors with a significance and 'reality' over and above the correlations from which they have been derived. It is tempting too to speculate about possible explanations, and to entertain concepts of factors as representing 'causal' entities in the mind of man. Such reification of factors, as Thomson and others have pointed out, is not to be encouraged. Any factor is defined by the tests from which it is derived, and it is best to think of its 'reality' simply in terms of its usefulness in the description and prediction of human behaviour.

Spearman's hypothesis that the correlations in his hierarchical matrix were due to a single 'factor' common to all the tests, while being a sufficient explanation was shown by Thomson (1916) not to be a necessary one. The conditions under which Spearman's hypothesis could be both sufficient and necessary have recently been re-examined by Hogben (1955) using a model of dice throwing in a game of chance.

In this game several players take part: they will be thought of as the tests. A number of umpires (representing the factors) also take part. Both players and umpires throw dice simultaneously and each throw can be taken as representing the performance of a child on the tests. After a throw each player adds to his own score some constant multiple of the score obtained by each umpire. If now the correlations between the players' scores are calculated the resulting matrix will have one common factor when only one umpire has taken part in the game, two common factors when two umpires have taken part, and so on. Hogben goes on to show that if the tetrad-differences are found to be zero with two or more umpires taking part then this can only occur when the composite score of the umpires is replaceable by that of any one alone.

To return to our earlier discussion, the idea that human behaviour could possibly be described in terms of a small number of factors was attractive. It was not new for it had already been embodied in the 'faculty' theories of earlier psychologists. The 'sampling theory' which Thomson used in his controversy with Spearman gives plausibility to the idea and, though the neuro-physiological concepts which inspired Thomson's thinking now need revision, the probability model on which it is based is still informative.

Thomson's sampling theory.

The name is a misnomer for this is not a theory in the usual sense of the word, but rather, an interpretation of how, under given assumptions, the mind may be imagined to function. Since the interpretation was put forward psychologists and statisticians alike have found it helpful and stimulating. Vernon (1950, p.31) acclaims it, and Bartlett (1953) states that, "sooner or later it will be necessary to consider mental activity in terms of the functioning of the brain, and Godfrey Thomson's interpretation may be regarded as a first attempt in this direction". The interpretation was put forward, as already mentioned, as an alternative explanation to Spearman's two-factor theory of 'hierarchical order' in matrices of inter-test correlation. Very briefly it is this - each test a person performs may be thought of as calling upon a sample of the 'bonds' which the mind can form. Tests differ in their complexity, some calling upon more 'bonds' than others for their performance. People also differ in their ability to do tests, and this difference - whether it be attributed to nature or nurture - is thought of as the ability to form the special 'bonds' to do the test concerned.

Interpreting human behaviour in test situations in this way Thomson was able to demonstrate that the laws of probability, or chance alone, are sufficient to cause a matrix of correlations between tests of mental ability to tend to show 'hierarchical' order. Indeed, in the special case where comprehensive tests of general mental ability in Spearman's sense are assumed, tests which in Thomson's terminology sample all the 'bonds' of the mind, the most probable outcome - statistically speaking - is a perfect 'hierarchical' matrix.

Important as this demonstration was, it did not end here. Spearman's two-factor theory, as already mentioned, soon had to be replaced by a multiple-factor theory providing for group factors as well as for a general factor and specifics. The sampling theory again provided an adequate explanation. To demonstrate the point a slight digression is necessary. An obvious criticism of the sampling theory as outlined above is that a high correlation between two tests would seem to imply that each had sampled almost all the bonds of the mind. Thomson had foreseen this and had made adequate provision for it in his theory by introducing the idea of 'sub-pools' of the mind. Two tests which correlate highly are thought of as sampling the same region of the mind. The situation is probably ~~like~~ MUCH AS Vernon pictures it (1950, p.31), when he states that "factors over and above g arise, partly perhaps from hereditary influences, but mainly because an individual's upbringing and education imposes a certain grouping on his bonds". It is worth adding too that whereas few psychologists to-day subscribe to a theory of strict localization of brain function there is no doubt that some functions have special associations with particular regions. Suggestive evidence on the point where human~~s~~ beings are concerned appears in a statistical review by the writer of a paper by Busch entitled "Psychical Symptoms in Neurosurgical Disease", (Maxwell, 1955). More convincing evidence from experiments with monkeys is reported by Delgado (1952), and by Rosvold, Mirsky and Pribram (1954). The latter found that monkeys which were dominantly aggressive in group situations before amgydalectomy tended to fall from top to bottom in the dominance hierarchy after the operation. The changes in behaviour were shown to be positively associated with the length of time the preoperative dominance relationships had existed. They were not related to differences in extent of the lesions but were consistent with differences in damage to the basolateral nuclei of the amygdala. The writers state,

Sholl

"it is probable, therefore, that one or more of the discrete structures in the temporal lobe are critical in bringing about the alterations in aggressiveness".

The Statistical Approach.

Returning to the factor concept the important and very striking thing is that if the sampling interpretation is assumed, then, as Bartlett has demonstrated (1937-38), statistical entities with all the properties of general, group and specific factors emerge and can be given a rigorous definition in terms of it. To quote him, (1953, p.31), "then for any individual the mean value of all his components (or at least all those which may be sampled by tests of some class under consideration) is defined as his 'general ability' g. A specific factor s is merely the contrast between the mean value of the components sampled by a particular test, and the average g of all components. If two tests each represent random samples from all the components, the specific factors s_1 and s_2 corresponding to the two tests will be independent. If they are not random samples, but are taken from a subset of the entire set of components, they will be correlated, and it will be necessary to introduce a further 'group factor' to account for this correlation".

What Thomson emphasises is that it is the laws of chance and not psychological laws which make the description of mental function in terms of a few factors conceivable. He was always careful to warn against the reification of factors. But if we get our interpretation of theoretical concepts clear the danger will not arise. Moreover, since Thomson wrote, psychologists - especially Tolman and Hull - have found it possible to speak of mental phenomena in a more helpful way, though the inferred phenomena themselves remain as 'unobservables'. The concept of 'intervening variable', has been introduced and used with some success. This concept does not differ qualitatively from the factor concept. Hull's (1943) habs, pavs, wats, and so on, immediately come to mind. The warning about employing such concepts is that they should be anchored firmly to identifiable and measureable quantities both at the stimulus and at the response end. At first sight factors may appear more elaborate concepts than habs: yet if one takes 'verbal ability' as an example, it is easy to convince oneself that it is such stuff as habs are made of.

The 'Bond' as Unit.

The unit Thomson chose for his theorising is of considerable interest in psychology. He did not try to define it rigorously. Attempted definition of the 'elements' of a science, as Hilbert has shown for the definition of a 'point' in geometry, only leads to vicious circles. Certainly Thomson puts our minds at ease by saying that on the mental side a 'bond' may be thought of as Thorndike did, as a 'connection' established when a habit is formed. And on the physical side he says we may think of it in terms of neuron arcs. But it is clear that sign-significate relations or stimulus-response connections would do just as well, depending on which camp one favours in the field of learning theory. Moreover, if one is physiologically inclined, 'bonds' may be thought of in terms of 'facility' at the synapses. Vernon (1955) has recently suggested that the term 'schema' - in the sense that Head and Bartlett use it - is preferable to 'bond', as being "less apt to suggest the mechanical operation of associations". This suggestion has much merit.

The need for some primary element, or basic unit, in psychology has long been felt. In contrast with physiology for which the nerve cell, or neuron, serves the purpose, psychology seems to require something implying a connection, or link. In retrospect it would appear that the notion that, 'associations' or 'bonds' are formed in the brain when mental events occur is the one which has persisted in psychological thought through the years. Its history is traced succinctly from Aristotle's 'processes' down to Pavlovian conditioning, by G. Humphrey in the first chapter of his book on "Thinking" (1951). A novel development of the idea has now been witnessed. Sir Ronald Fisher (1953, p.1) has recently suggested that "Statistical Science was the peculiar aspect of human progress which gave the twentieth century its special character". In the statistical treatment by Thomson and Bartlett of the sampling theory using the 'bond' as unit, the 'adventure of ideas' has come full circle, and the outcome as we have seen is striking.

The fundamental factor equation and its solution.

Parallel with the above development multiple factor ideas already referred to were taking shape largely following Thurstone's generalization of Spearman's theory - but always capable of interpretation in terms of the sampling theory - the fundamental factor equation being stated in terms of,

1. factors common to two or more tests,
2. factors specific to the separate tests, plus 'error'.

The latter assumption which causes the number of factors to exceed the number of tests gives to the factor equation a mathematical indeterminancy which has not been resolved to everybody's satisfaction. Viewed psychologically in the light of the sampling theory, where the 'specific' for any test is thought of merely as a contrast between the mean value of the components sampled by that test and the average of all components, this indeterminancy would appear unimportant. Either Thomson's (1951, pp.221 et seq.,) or Bartlett's (1937-8) estimation equations - the two converge as the number of tests increase - would give the experimenter all the information he could justifiably expect from his analysis.

The fact, however, that the fundamental factor equation does not make specific allowance for 'error' is one to which the statistician may rightly object. This deficiency cannot be passed over lightly by the psychologist for the whole reliability of his work is at stake. While he continues to employ his present models replication of factorial studies to establish the reliability and invariance of any factors he claims is obligatory. Meanwhile the statistician must not bask in the sun of self-rightousness for to paraphrase Professor M.G. Kendall's words (1951, p.18) - the statistician working in other fields can hardly afford to throw stones at the factor analyst if the latter does not feel able to include an error term in his equation and make an assumption about its normality, independence and equal variance for each of his tests; the statistician deals with the error problem only by 'assuming it away' in terms of these assumptions, without always justifying the procedure.

Choice of a suitable factor model.

The principal component model we have already considered: when used with unities in the diagonal cells it is not compatible with the psychologist's needs. At any rate it has statistical shortcomings (Maxwell, 1956) which its adherents seem to play down, (Wrigley and Neuhaus, 1955). In the first place the components obtained by it are not invariant under changes of scale; this means that values found if the covariance matrix instead of the correlation matrix is employed are not proportional. As Thomson (1951, chapter 21) has pointed out such proportionality is an essential requirement of any satisfactory factor method. Secondly, little if any work has been done on the sampling issues involved in principal component analysis. Thirdly, in order to test the significance of successive latent roots, either all the roots must be found, or the value of the determinant of the matrix must be calculated. Both calculations are laborious. The available test too (Bartlett, 1950) is not a test of the number of roots which differ significantly from zero. Rather is it a test of whether after one or more roots have been taken the remaining roots differ en bloc from these and individually from each other.

If then disfavour is thrown on the principal component model what is left? By far the most common factor method in use is the centroid, or simple summation, method. It is an approximate method only. For nearly a quarter of a century now it has resisted all attempts at a rigorous mathematical and statistical treatment. This has been unfortunate and has contributed a good deal to throwing factor analysis into disrepute among statisticians and psychologists alike. In justice to it, however, we hope to show that with careful handling it probably gives sufficiently accurate results for most practical purposes. *

---

* Since this paragraph was written a paper entitled "A statistical examination of the Centroid Method", by D.N. Lawley (1955) has appeared. The synopsis of it reads as follows -

> A modified form of the centroid method used in factor analysis is described. Various large sample results are obtained, including a test of the significance of the residuals. The method is compared with the corresponding form of the maximum likelihood estimation and its efficiency is investigated. A numerical illustration is given of some of the foregoing theory.

Lawley finds the efficiency of the centroid method to be much higher than was previously suspected - a finding which gives concrete support to the suggestions made at the end of the paragraph in question.

Apart then from the centroid method and one or two other approximate techniques, which are not discussed here there remains the maximum likelihood approach to the factor problem by Lawley (1940, 1955).

Earlier doubts (Kendall, 1950) about the validity of his Method I (1940) - except where the number of tests is very small - have not been confirmed. Furthermore general doubts (Kendall, 1955) about any use of maximum likelihood methods for fitting factors seems to reveal a lack of understanding of the factor problem. We may conclude then, as Bartlett does, that Lawley's method "should, as far as we know at present, give a satisfactory solution of the problem", (1953, pp.27-8). The method too seems to possess more of the optimal properties of a perfect method than either the principal component or the centroid approach provides. The results it gives are invariant under change of scale in the variables, consequently loss of information arising from any use of 'standardized scores' can be obviated. It provides too, for samples even moderately large, a satisfactory test of the number of significant factors. The sampling theory of the method also is now well in hand. This offers far-reaching prospects for it will make possible a comparison of results obtained from different experiments. Like the principal component method, the maximum likelihood method also will be a practical proposition as electronic computing facilities become available and it would appear, on statistical as well as on psychological grounds to merit serious consideration.

A precise statement of the factor problem and a definition of a factor.

So far in this chapter remarks about factor analysis and about factors have been of a general and imprecise nature. This has not been altogether unintentional for it is in keeping with the method of exposition found in most of the literature on the subject. However, now that the die has been cast in favour of the maximum likelihood approach it will be in the interests of clarity to state the problem - to which Lawley has given a satisfactory solution - in more precise terms.

Quoting Howe (1955, p.1) - "the usual method of stating the problem is: can a p-dimensional complex of random variables be represented adequately by $m < p$ variables? Obviously 'adequately' is the ambiguous word which causes the different interpretations. A linear relationship is then assumed and the model is expressed.

$$x_{ij} = \sum_{k=1}^{m} a_{ik} y_{kj} + e_{ij}, \qquad \begin{array}{l} i = 1,\dots p, \\ j = 1,\dots N, \end{array}$$

where N = sample size,

x_{ij} = the value of the random variable x_i for the j-th person,

y_{kj} = the value of the k-th common factor for the j-th person,

a_{ik} = the factor loading of the i-th random variable for the k-th common factor; and e_{ij} is a random error such that

$E(e_{ij}) = 0 = E(e_{ij}e_{nj}) = E(e_{ij}y_{nj}), \quad i \neq n$ ....".

Lawley in his solution makes the assumption that the y_{kj} are random normal variates with variance one and mean zero.

A factor may ~~now~~ be defined statistically as a variate which is a linear function of a set of observed variates. It is itself non-observable and so incapable of direct verification. On the psychological side - to use Thomson's language - it may conveniently be thought of as an imaginary test. Its nature can only be inferred from antecedent knowledge of the tests which contribute to it, due consideration being given to the weights allotted to the latter in the linear equation.

Summary.

In this chapter an effort has been made to show that the motivating force behind the development of factor analysis came from an interaction between mathematical and psychological ideas.

Having obtained a correlation matrix it was natural to try to describe it. The first effort to do this resulted in Pearson recommending that the matrix should be transformed into orthogonal components or axes. Later, with regard to correlation matrices derived from cognitive variables, Spearman observed a tendency towards 'hierarchical order' and the possibility of accounting for the correlations by a single factor. Thomson pointed out that Spearman's idea while offering a sufficient explanation of the hierarchical tendency was not a necessary explanation of it. A controversy ensued and in the process Thomson put forward his 'Theory of Bonds', as a possible explanation of the way the mind might be supposed to function. The statistical ideas behind this theory were further developed by Bartlett.

Spearman's ideas were later generalized by Thurstone to meet cases where - though the correlation matrices were not hierarchical - the correlations could be accounted for by postulating a small number of factors. Thurstone too developed simple summation or what he called the Centroid Method for extracting factors. This method, it was pointed out, bore many resemblances to Pearson's principal axes method but had an affinity with Spearman's line of thought in that it aimed at keeping the number of postulated common factors as small as possible.

Psychological objections to the principal axes method and statistical objections to the centroid method were then noted, and the maximum likelihood approach to the factor problem by Lawley, which simultaneously satisfies the psychologist's ideas about the function of factor analysis and meets the dictates of sound statistical theory, was then recommended. In the light of this recommendation a concise statistical statement of the factor problem was then made and a psychological factor defined.

Chapter III

THE DATA.

In this chapter an account is given of the data on which the discussion in later chapters is based. The samples and tests are described. Details of the extraction of factors by the centroid, the principal component and the maximum likelihood methods are given, and, in the case of the latter two methods, tests of significance of the factors are employed.

The Mr. James Sample.

The first sample consists of 810 children, in the age range 11-14 years, drawn from London secondary modern schools. The writer is indebted to Mr. David James for permission to use these data. For convenience the sample will be referred to as "Mr. James's sample". He administered a battery of twelve tests. The first five variables are from Thurstone's Primary Mental Abilities (1939). The "11-17" version was used and the variables concerned are -

1.	Verbal ability	(0.92)
2.	Spatial ability	(0.96)
3.	Reasoning	(0.93)
4.	Number	(0.89)
5.	Verbal fluency	(0.90)

The numbers in brackets given after each test are the reliability co-efficients of the tests as quoted by Thurstone.

The remaining seven variables are measures of neurotic tendencies. They are

6.	Questionnaire	(0.74)	(0.89)
7.	Ways to be different	(0.88)	(----)
8.	Word likes and dislikes	(----)	(----)
9.	Worries and anxieties	(0.76)	(----)
10.	Interests	(0.86)	(0.93)
11.	Annoyances	(0.81)	(0.85)
12.	Sentence completion test	(----)	(0.59)

Copies of the versions of these inventories are given in Appendix A. They were specially devised for this and similar studies with children, by members of the staff of the Psychology Department at the Maudsley Hospital. The 'questionnaire' is based on the Maudsley Medical Questionnaire (Eysenck, 1947), while in the compilation of the other inventories relevant research work by Bennett (1945), Himmelweit and Petrie (1951) and earlier writers was helpful.

The reliability figures quoted in the first column after the neurotic inventories are 'split-half' reliability co-efficients corrected for length, given by Thorpe (1953). They are based on the answer-patterns of 10^{0} children chosen at random from a sample of over 900. The figures in the second column are also corrected 'split-half' reliability co-efficients.

That for the questionnaire is given by Eysenck (1947) and is based on a large sample of neurotic male patients. The others were calculated by the writer and are based on a sample of 104 adult male controls. In general these reliability co-efficients - while not as high as one would like - are reasonably good for inventories of this kind.

The matrix of product-moment correlation co-efficients between the twelve variables is given in Table III. This matrix was factorized by the centroid, Lawley's maximum likelihood, and Hotelling's principal component method. In the latter case the first four latent roots of the matrix only were obtained.

The Centroid Analysis.

When finding the centroid loadings an interesting situation arose. The fact that the algebraic sum of each column is positive (see the last row of Table III) allows one to proceed immediately to extract the first centroid, if the rules of thumb given by Thurstone (1947, p. 165) are followed. There he states that if all the column sums are positive "no reflections are necessary". It only remains to decide what is to go into each diagonal cell and proceed to calculate the first factor loadings. The first residual matrix can then be obtained, signs reflected following Thurstone's criterion and a second factor extracted, filling the diagonal cells as before.

The results are given in Table IVa. There it is seen that the amount of variance extracted by the first factor (last row of the table) is 9.85% while that extracted by the second factor is 26.18%. This at first appears unusual. With Hotelling's principal components method the above results would clearly be impossible for the first axis is the major axis of the 'ellipsoidal' test space and automatically extracts maximum variance.

The conditions under which maximum variance will be extracted by a centroid factor are, to the writer's knowledge, nowhere made clear. In practice the tendency is to make the algebraic grand total (T) of the entries in the matrix, whether it be a correlation or a residual matrix (but omitting the diagonal cells), a maximum. This is achieved by sign reflection. With a large matrix the realisation of this end can be difficult and tedious. Since Thurstone's method of factor analysis is complete only after rotation to simple structure it is not essential for his work to ensure the extraction of maximum variance with succeeding factors. But many factor analysts do not hold with the 'simple structure' idea, and for them it is desirable to extract maximum variance with each factor .

At present there are at least two approaches to the sign-reflection problem which help, the method described by Thomson (1951,p.71), and Holley's criterion(1947). The former can be used instead of Thurstone's approach, the latter is most valuable in conjunction with it. If the method described by Thomson - the aim of which is to reflect those variables which will minimise the number of negative signs in the correlation matrix and so tend to maximise T - is adopted, the signs of variables 6, 7, 9, 10 and 11 will require to be reversed.

Holley considers the effect on T of reflecting variables two at a time. For any two variables he is able to specify the conditions under which reflection will contribute a positive increment to T. Omitting the argument behind his method, which is lengthy, the procedure (Holley, 1947, p.265) is as follows -

1. find the variable with the lowest column sum, not including diagonal values,
2. in this column and row, find the coefficients of correlation which have a higher positive value than the column sum.
3. compare these pair combinations, and
4. reflect the two variables when the difference between twice the correlation and the sum of the two column sums, excluding diagonal values, is both positive and maximum,
5. if the criterion is not met in the case of the variable with the lowest column sum consider the variable with the next lowest sum, etc.

Holley's criterion when applied to the above matrix indicates reflection of the same variables as did the method described by Thomson.

The analysis was now continued using Thurstone's method supplemented by Holley's criterion. Five centroids in all were extracted, on each occasion the highest entry in a column was placed in its diagonal cell. The results are entered in Table IVb. No iterations were done at this state. The communalities (h^2) which result from the factor loadings obtained are given in column 6 of the table. In column 7 the highest r for each column, which was used as communality when getting the first factor loadings, is entered, and in column 8 the absolute sums of the rows (or columns) of the correlation matrix, omitting diagonal cells, are given. An examination of the entries in column 6, 7 and 8 gives support to the claim that Thurstone's method of filling the diagonal cells by the biggest r in each column tends to underestimate the communalities of tests which correlate relatively highly with the other tests in the battery, and to overestimate the communalities of tests which correlate less well with the other tests.

The Maximum Likelihood Analysis.

Next the correlation matrix in Table III was factorised by the maximum likelihood method. An hypothesis that three factors would be sufficient to account for the correlations was made and the loadings of the first three factors of the centroid method were used as guesses of the maximum likelihood loadings.

After four iterations good convergence to three decimal places was reached. The loadings for the third and fourth iterations are given in Table Va.

A residual matrix was now calculated and tested for significance by the formula,

$$\chi^2 = (N - 1)\, S r_{ij}^2 / s_i^2 s_j^2, \quad i < j, \quad (i, j = 1, 2, \ldots\ldots p)$$

where r_{ij} is the residual in the i-th row and j-th column, s_i^2 is the specific variance of the i-th test, and N is the number in the sample. The degrees of freedom are $\frac{1}{2}(p - k)^2 - \frac{1}{2}(p + k)$, where p represents the number of tests and k the number of factors (Lawley, 1940).

The residuals were found to be significant and so the original hypothesis of three factors had to be discarded. An hypothesis of four factors was now set up, the loadings of the fourth centroid factor being taken as first guesses for the fourth maximum likelihood factor. One iteration only was performed and the residuals when tested again were found to be significant. A fifth factor was now fitted again using the fifth centroid loadings as a first guess. Two complete iterations of the five factors were performed at this stage. The results are given in table Vb. A residual matrix was again found and when tested gave a value of chi-square equal to 63.307

with 60 degrees of freedom which is still highly significant. No further factors were extracted at this stage since the amount of variance which they might be expected to account for would be negligible: that for factor V was only 2.2%.

The fact that as many as five factors are significant is due doubtless to the relatively large size of the sample (N = 810). Had the sample only an N of 300, chi-square would then have been 23.45 and would have been well below the 5% significance level with 16 degrees of freedom.

Here it is well to repeat that statistical significance is not synomymous with psychological meaning, and judging by the very small amounts of variance accounted for by factors 4 and 5 (table Vb) it is doubtful if they are of any value from a psychological point of view. It is commendable, however, to err on the side of having more significant factors than one can interpret ~~than vice-versa~~: while significance is not a sufficient criterion of a psychological factor it is a necessary one.

The formula for chi-square used above is an approximate formula for use with large samples given by Lawley in his original article (1939). The exact formula he gives there as -

$$\chi^2 = n.\ln|C|/|A|$$

where ln stands for natural logarithm. A is the observed correlation matrix with $n = N - 1$ degrees of freedom, and C is the fitted correlation matrix given by $C = LL' + V$ where L is the matrix of factor loadings and L' its transpose, and V is the diagonal matrix of specific variances.

Later Bartlett (1950) suggested that for moderately large samples a better multiplier than n may be

$$n' = n - \frac{(2p + 5)}{6} - \frac{2k}{3}$$

but this has not been rigorously demonstrated. As before p is the number of variables, and k is the number of factors.

In a recent article Lawley and Swanson (1954) point out that the use of the exact formula, unless iteration has been carried to the stage where good convergence of the factor loadings has been attained, may give results far worse than the approximate method. They proceed to show how the difficulty can be overcome by the use of a slightly more complicated expression.

But it would appear that Lawley's approximate formula is more accurate than he himself has emphasised. On the theoretical side this is pointed out by Bartlett (1950, p. 84). As an empirical check the writer calculated a chi-square value by the approximate formula for the residuals after two factors for the complete sample of 400 in Lawley and Swanson's article. He obtained a value of 8.7, which is exactly that got by the writers using the more elaborate procedure to which reference has been made. It does not follow, of course, that for smaller samples agreement would be so good.

Henrysson (1950) too has tested Lawley's approximate formula. Using artificial data he chose nine variables arranged to have only one common factor, and took 12 samples each of size 200. Twelve correlation matrices were calculated and one factor fitted to each by the maximum likelihood method. None of the residual matrices when tested by the approximate formula gave a significant value of chi-square, a result which is in full agreement with the known structure of his matrices.

The Principal Component Analysis

Finally, Mr. James's correlation matrix was subjected to an Hotelling principal component analysis. The first four factors only were obtained, the four latent roots in order of magnitude being 3.710, 1.642, 1.272 and 0.966. The amount of variance accounted for by these factors is then 30.9%, 13.7%, 10.6% and 8.5%, respectively. The factor loadings are given in table VI.

The Significance of Principal Components.

In the example just described no effort was made to test the significance of the latent roots obtained in the principal component analysis. The appropriate test is one by Bartlett (1950), but it has the disadvantage that before it can be applied either all the latent roots of the correlation matrix must be known, or the value of the determinent of the matrix found. In either case the calculations are laborious, even for a 12 x 12 matrix.

To illustrate Bartlett's test an example from an article by Emmett (1949) in which only nine variables are concerned was chosen. Emmett derived his 9 x 9 correlation matrix from a 17 x 17 matrix reported by P. Slater (1943). The variables concerned included seven 'spacial', four 'non-verbal' and six 'verbal' tests, and had been administered to 211 boys and girls, aged 11 . Slater in a centroid analysis had failed to find evidence of a 'spacial' factor, but Emmett employed Lawley's maximum likelihood method and found three factors to be significant - the third at the 0.05 level, one spacial test having a highly significant loading on it.

The correlation matrix used by Emmett together with the loading of the three factors, both before and after rotation, are given for reference in Table VIIa.

The writer extracted four 'factors' from Emmett's correlation matrix by Hotelling's principal component method. The four latent roots were 4.677, 1.262, 0.840 and 0.556 and the corresponding factors accounted for 51.9%. 14.0%. 9.3% and 6.2% of the variance respectively. The loadings are given in table VIIb, adjacent to those found by Emmett.

In order to apply Bartlett's test it was now necessary to calculate the value of the determinant of the correlation matrix in table VIIa. Since only the first four of the latent roots of the matrix were known it was not possible to use the formula

$$|R| = L_1 x L_2 x \ldots\ldots\ldots\ldots L_9$$

where the L's stand for the latent roots.

The usual algebra text-book methods for evaluating a determinant are of little use in practice as they easily get out of hand. The method here employed was Aitken's pivotal condensation procedure. The calculations are similar to those for finding the inverse of a matrix (Thomson, 1951, pp. 65 and 205). The value of the determinant of the matrix is given by the product of the first (n - 1) factors extracted when reducing the pivots to unity.

The procedure gave the value

$|R| = 0.0088762.$

Bartlett's test, as he himself states (1950, p.78), "is formulated to indicate the significance of the residual roots", after the removal of the largest roots. The formula is

$$^2 = - n - \frac{(2p + 5)}{6} - \frac{2k}{3} \ln R_{p-k}$$

with $\frac{1}{2}(p - k)(p - k - 1)$ degrees of freedom, after k roots have been determined, where

$$R_{p-k} = |R| / (L_1 x L_2 x \ldots L_k) \left(\frac{p - L_1 - L_2 - \ldots - L_k}{p - k}\right)^{p-k}$$

In this formula L_1 stands for the latent roots, n is equal to (r - 1) where r is the size of the sample, and p is the number of variables.

Bartlett continues by pointing out that it is often convenient to present a complete chi-square table, but since the multiplying factor in the formula for chi-square given above changes with k the complete table is not strictly additive unless a constant factor is substituted. The constant factor he suggests is $C = -(n - p + \frac{1}{2})$.

The latter approach is the one used here. The constant factor in our case is -201.5 and the contribution of each latent root to the total chi-square is given below. (Table VIII)

In a later article Bartlett (1951, p.1) points out that it is more appropriate to calculate the degrees of freedom for roots after the first one or so by the expression

$$\frac{1}{2}(p - k - 1)\ (p - k + 2).$$

In our case the degrees of freedom for the first latent root are those found by the earlier formula, and the degrees of freedom for roots two, three and four by the latter. The significance of all four roots is clear from the table VIII.

The best approximation for the total chi-square, Bartlett tells us, would be

$$-n\left[n - \tfrac{1}{2}(2p + 5)\right]\ln|R|$$

In our case this is equal to 974.03. This would mean that the contribution of the remaining five roots to the total chi-square is 71.11 with 16 degrees of freedom. It is probable then that the fifth latent root is also significant. The degrees of freedom for it are 9, and a value of chi-square equal to 16.92 would be significant at the 0.05 level.

Mr. Nigniewitsky's sample.

The next matrix we will consider gives the correlations between nine variables selected from a study by Mr. R. Nigniewitsky. In 1954 he investigated the opinions and attitudes of the major French political parties. The sample consisted of 380 adults containing about equal numbers of males and females. Each person answered an attitude inventory consisting of 396 items. These items covered, amongst others, variables entitled,

1. Rigidity
2. Intolerance of ambiguity
3. Sandford Gough rigidity scale
4. Rhathymia
5. Ascendence
6. General activity
7. Neuroticism
8. Psychopathic attitude
9. Neurotic attitude.

The matrix of intercorrelations is given in Table IXa. Its special interest from our point of view lies in the well defined and easily recognisable patterns amongst the correlation

The matrix was factorized by the centroid method only, but to check on the accuracy of the results one iteration was performed. Three factors were extracted. In the first instance, following Thurstone, the highest r in each column was used in the diagonal cell of that column. This procedure was repeated in the first and second residual matrices. However, in the iteration of the factors which followed the communalities obtained from the first estimates of the factor loadings were employed.

The factor loadings and the estimates of the communalities obtained in each iteration are given in table IXb. It is worth noting here how little the communalities changed with the iteration. This confirms the view that although Thurstone's 'Highest r' procedure - as we have seen in an earlier example - is not the best procedure to adopt, the communalities obtained by it, even without iteration, are themselves very good estimates of the true communalities. It follows too that one iteration in a centroid analysis may be expected to give a reasonably accurate picture of the true centroid loadings.

Replicated experiments.

The need for replication of experiments in factorial work already has been emphasised. It is only by replication that information about the 'stability' of the psychologist's factors can be found. The examples now to be presented were chosen to provide a basis for a discussion of 'stability', and to supply data on which 'loadings' obtained in replicate experiments could be compared.

The data were collected by members of a research unit at the Maudsley Hospital and the writer is indebted to Professor H.J. Eysenck for permission to use them.

A battery of tests, from which the list below is a selection, was administered in 1950 to a sample of 104 children, aged 11 and reported by their teachers to be 'neurotic'. In the following year the same battery of tests was administered to a similar group of 148 children attending the Maudsley child guidance clinic.

The variables for which data are reported here are,

Number.	Name of Variable.
1	Annoyance
2	Ways to be different
3	Neuroticism questionnaire
4	Worries and anxieties
5	Interests
6	Verbal ability
7	Spacial ability
8	Reasoning
9	Number
10	Verbal fluency
11	Manual dexterity
12	Track tracer
13	Tapping
14	Reaction time (1)
15	Reaction time (2)

The first ten variables have already been discussed in Mr. James's sample: the remaining five are well known psychometric measures, detailed description of which need not detain us here.

Using different combinations of these variables three correlation matrices for each of the years 1950 and 1951 were chosen. This gave six correlation matrices in all. The names allotted to each and the variables included in them are given in the following table -

Name	Table	Variables Included
1950A	X	1 to 6, 8 and 10
1951A	XI	ditto
1950B	XII	1 to 10
1951B	XIII	ditto
1950C	XIV	1 to 6, 8 and 10, 11 to 15
1951C	XV	ditto

The first two matrices were factorised by both the centroid and the maximum likelihood methods. The results appear in tables X and XIa and b respectively where full details of the convergence of the factor loadings with successive iterations are given.

In the 1950A sample good estimates of the communalities of the tests were known from the parent study from which the tests were taken, and these were used in getting the first estimates of the centroid loadings. For the iteration, the communalities found at the previous stage of the analysis were always used. Reasonably good estimates of the maximum likelihood loadings were also known from the parent study. They are shown in columns A and B of table X. But in spite of the initial advantage of having prior knowledge of the loadings, inspection of successive iterations in the maximum likelihood procedure shows either that convergence tends to be fairly slow or that, in this particular instance, without more decimal points greater stability cannot be achieved. In this respect ready access to electronic computing facilities would be very advantageous.

In the 1951A sample another case of the problem met with in "Mr. James's sample" is seen. Thurstone's (1947, p.165) rules of thumb again fail to extract maximum variance with succeeding factors. This led to considerable inconvenience, for - neglecting to check the point - the writer launched out on a maximum likelihood analysis of the 1951A correlation matrix using the iterated centroid loadings as initial estimates. As is clear from the data in table XIb there was no sign of stability being reached after four iterations. To hasten the process new guesses of the loadings, consistent with the direction in which convergence was tending, were made. Further iterations were carried out but again with disheartening results. At this stage the defect in the centroid analysis was detected. Holley's criterion was applied, new centroid loadings (table XIa) were calculated, the first extracting 29% of the variance and the second 17.7% as compared with 18.8% on the first and 28.7% on the second in the earlier analysis. The revised centroid loadings were now used in the maximum likelihood procedure and (table XIb) after seven iterations good convergence was attained.

Inspection of table XIb shows that had the original maximum likelihood analysis been continued the desired results would eventually have been achieved. The amount of calculations involved, however, would have been phenomenal.

The residuals of the 1950A and 1951A matrices, after two factors had been fitted, were tested for significance by Lawley's approximate formula. The former was just significant at the 0.05 level, the latter was not significant: the values of chi-square with their degrees of freedom were -

Sample	N	Chi-square	d.f.	significance
1950A	104	24.16	13	0.05
1951A	148	20.28	13	N.S.

Since the second factor in the 1951A sample accounts for as much as 17.7% of the total variance and since the residuals after two factors are almost significant, the significance of this factor may be considered established.

In the 1950A residual matrix the two highest residuals were 0.108 between tests 1 and 2 and -0.079 between tests

4 and 10. However, since the residuals were only just significant at the 0.05 level no attempt was made to calculate a third factor.

The 1950B and the 1951B correlation matrices differ from the 1950A and the 1951A matrices only by having two additional variables, numbers 7 and 9 being included. But the correlations between them and the variables 6 to 10 differed very considerably in the two samples. The relevant details are -

1950

Tests	1	2	3	4	5	6	7	8	9	10
7	-137	-095	-096	007	-238	142	...	290	-066	-004
9	094	-075	-134	-046	-072	164	-066	218	...	210

1951

Tests	1	2	3	4	5	6	7	8	9	10
7	-162	-136	017	-010	-216	533	...	419	282	276
9	023	-154	-118	-096	-096	532	282	532	...	518

To investigate the effect of these rather big discrepancies on the factor loadings already found for the 1950A and 1951A matrices maximum likelihood analyses of the 1950B and 1951B matrices were carried out. The results are shown in tables XII and XIII.

In the case of the 1950B matrix an hypothesis that two factors would be sufficient to account for the correlations was made and two factors were fitted. The residuals were then tested and found to be significant at the 0.02 level, chi-square being equal to 43.29 with 26 degrees of freedom. The inclusion of the two additional variables - though the intercorrelations with the other variables were relatively small - had thus raised the level of significance of the residuals from the 0.05 to the 0.02 level.

For the 1951B matrix - in view of the relatively large correlations added by including variables 7 and 9, and the fact that the sample size here is 148, as compared with 104 in the 1950 sample - it was decided to postulate three factors. This proved to be justified for a test of significance of the residuals after two factors with five iterations had been removed was significant well beyond the 0.01 level.

The third factor extracted 4.6% of the total variance, tests 7, 3 and 4 having the highest positive and tests, 5, 9 and 1 the highest negative loadings on it.

In view of the relatively small size of the sample (N = 148) no attempt was made to extract any further factors nor were the risiduals after three factors tested for significance.

The Signs of Factor Loadings

The fitting of the factors to the 1951B matrix revealed a new point of interest. Tables X and XII show that for the 1950A, and 1950B and 1951A matrices the first maximum likelihood factor fitted in each case has positive loadings on variables 1 to 5 and negative loadings on the remaining variables. But for the 1951B matrix - table XIII - the signs of the factor loadings are reversed, 1 to 5 being negative positive and 6 to 10 inclusive being positive negative. The reason for this can be seen by looking at the following figures -

1950A		1950B	
7.234	-3.676	7.234	-4.468
-3.676	2.946	-4.468	4.854

1951A		1951B	
7.520	-1.718	7.520	-2.664
-1.718	3.412	-2.664	9.594

In each case the original correlation matrix was partitioned into four quadrants to separate the neurotic variables (the first five), from the cognitive variables. The algebraic sum of the correlations in each quadrant (omitting diagonal cells) was then found. These are the figures given above. Looking at the entries about the diagonals only it is seen that for the first three samples the sums of the inter-correlations between the neurotic variables are larger than those between the cognitive variables. But for the 1951B matrix the situation is reversed. In it the sum of inter-correlations for the cognitive variables outweighs that for the neurotic variables with the result that the former take precedence as regards <u>sign</u> when the first factor is fitted.

When comparing the first factor of the 1951B matrix with the first factor of each of the other matrices it is of course legitimate to reverse the signs of its loadings. This interferes in no way with the 'goodness' of the factor 'fit'.

Before leaving these correlation matrices a further point may be noted. It concerns the entries in the N.E. quadrant of the partitioned matrices. These entries for the 1950 sample, viz. -3.676 and -4.468, are considerably greater than the corresponding entries -1.718 and -2.664 for the 1951 sample. The explanation suggested for this in the parent study was that the teachers' notion of a neurotic child was associated to a considerable extent with retardation on the scholastic side. It is not without significance too that this retardation was most seriously reflected in the correlations of variables 7 and 9 viz. Spatial ability and Number, with the remaining cognitive variables, Verbal ability, Reasoning and Verbal fluency.

Finally, matrices 1950C and 1951C which included all the variables, ~~expect~~ except number 7 and 9, were analysed by the centroid method and the results after one iteration are shown in tables XIV and XV. In each case three centroids were extracted. In the 1951C matrix a new situation was met. Thurstone's rules of thumb for sign reflection gave a first centroid accounting for 14.0% of the variance of the tests while the second centroid accounted for 19.6%. Here, however, Holley's criterion as a supplement to Thurstone's rules, would not work. To overcome the difficulty Holley's criterion was applied first. On its own it was found inadequate for getting all the column sums of the matrix positive (Thurstone, 1947) and it in turn had to be supplemented by Thurstone's rules.

Summary.

In this chapter the correlation matrices on which the discussion in later chapters is largely based are given. Passing reference is made to the variables employed and the samples tested, but in the main the chapter deals with the extraction of factors and the use of valid tests of significance.

The factorial methods employed are:

1. The centroid method
2. The principal component method
3. The maximum likelihood method

The correlation matrices include,

Mr. James's 12x12 matrix (N is 810)

Mr. Emmett's 9x9 matrix (N is 211)

Mr. Nigniewitsky's 9x9 matrix (N is 380)

The 1950A 8x8 matrix (N is 104) with its replicate the 1951A matrix (N is 148).

The 1950B, 10x10 matrix (which includes the eight variables from the 1950A matrix), with its replicate the 1951B matrix.

The 1950C 13x13 matrix (which also includes the eight variables from the 1950A matrix), with its replicate the 1951C matrix.

With regard to the centroid method of extracting factors it was noted that Thurstone's 'rules of thumb' for sign reflection did not necessarily lead to a unique solution. When used in conjunction with Holley's criterion however it was possible to achieve the desired effect of extracting maximum variance with succeeding factors - Mr. James's matrix and the 1951A and 1951C matrices afford examples.

As a consequence of this it was shown that the use of centroid loadings - which did not extract maximum variance with succeeding factors - as first guesses for the maximum likelihood loadings, might inadvertently lead to much extra labour. Table XIb admirably illustrates the point. This difficulty is magnified too by the slow rate of convergence of the loadings in the iterative procedure at present available for the maximum likelihood method (tables X to XIII). In view of this the prospect of electronic computing facilities soon being more easily available is welcomed.

The efficiency and ease of calculation of Lawley's approximate formula for testing the significance of residuals after fitting maximum likelihood factors was emphasised, and illustrated by reference to articles by Bartlett, Lawley and Swanson, and by Henrysson. This is in contrast to the laborious calculations involved in Bartlett's test for principal components. The latter test was demonstrated by testing the first four latent roots of Mr. Emmett's 9x9 matrix, but was avoided in the case of the first four latent roots of Mr. James's 12x12 matrix, where the calculations would have been almost double those for the smaller matrix.

Matters Pending.

Empirical rules for deciding on the significance of centroid factors are too numerous to be dealt with here and are allotted a separate chapter. Comparison of the loadings found by factorizing a matrix by more than one method is also postponed pending a discussion of the standard errors of factor loadings.

Chapter IV

HOW MANY FACTORS?

In the previous chapter no attempt was made to test the significance of factors found by the centroid method. In this chapter a few of the empirical tests generally employed will be reviewed and their validity checked against the results obtained in the maximum likelihood analyses.

The Centroid Method and Sampling Theory.

It was unfortunate for the factor analytic movement that the centroid, or simple summation, method of multiple factor analysis, which supplanted Spearman's two-factor approach, was well-nigh devoid of sampling theory. Sampling issues are barely mentioned in Thurstone's book. The result is that a factor model which (until Lawley's recent work, 1955) had resisted all efforts at rigorous mathematical and statistical treatment, has been the model most popular amongst psychologists. Consequently, factorial studies using the centroid method - no matter how high the reputation of the author, and no matter how much care has gone into the work - could always be held up to ridicule by the unsympathetic critic. The latter generally choses to overlook the fact that statistical significance and psychological meaning are not synonymous. Seldom can the analyst interpret with conviction as many factors as even the most cautious dictates would allow him extract. However, in the absence of valid tests of significance the degree of caution exercised unfortunately was free to vary and did so to a considerable extent.

It is often said that factorists in the British tradition tend to be, if anything, too stringent in the number of factors they extract while American factorists tend to err in the opposite direction. One has not to look far in the literature for evidence to support this statement. Professor P.E. Vernon, who in his privately circulated manuscript "How many factors?" was one of the first writers to draw attention to this contrast, supports his remarks by reference to a study by Woodrow (1950) who extracts ten factors with a sample of only 110 students who had taken 52 tests, and interprets six of these after rotation. In contrast to Woodrow, Vernon quotes Burt (1940) who recommends that a sample of 250 is needed for three factors to be significant, or a thousand for four factors. More examples, since Professor Vernon wrote his manuscript, are not difficult to find. In the cognitive field Taylor (1947) in "A factorial study of fluency in writing", using 10 tests and a sample of only 181, extracts 10 centroid factors and interprets eight of them. By contrast Moursy (1952) in his much quoted study entitled "The hierarchical organisation of cognitive levels", using 20 tests and a population of 166 boys extracts only 5 factors of which he says only three are statistically significant. In the personality field Wittenborn (1951) using 55 rating scales and a population of 140 patients extracts and identifies 7 factors while Trouton and Maxwell (1955) in a factor analytic study of 45 items from the Maudsley Hospital Item Sheet extract only 6 factors with a sample of 819 patients, and feel justified in interpreting only the first two with conviction, and two more tentatively. Many more examples could be given: two gross cases are that of Cattell (1947), and Wilson et al. (1954). Cattell in a study entitled "Confirmation and clarification of primary personality factors" with 36 variables and a sample of only 133 found justification in Tucker's criterion for extracting 13 factors - the last two to allow himself elbow space for rotation! Lastly, Wilson, Guilford, Christensen and Lewis (1954) in "A factor-analytic study of creative-thinking abilities", use 53 tests designed to measure aspects of creative thinking and with a population of 410 air cadets extract 16 factors. After rotation they proceed to identify 14 of them.

Empirical Tests of Significance.

It must not be suggested, however, that psychologists have carried on using the Centroid method of analysis quite oblivious of its statistical defects. The contrary has been the case. An overwhelmingly large number of 'logically' derived or empirically tested rules purporting to help the investigator to decide when to stop factoring have appeared in the literature. In the manuscript referred to above Vernon evaluates some two dozen of the 'indices of significance', as he calls them for use with the Simple Summation Method. His criterion of their validity was the very practical one of testing them on the results of two factor analyses carried out on two large samples where the tests used were well known to him and where the identity of the factors obtained was clear. As a result of his investigation he made certain recommendations. These are summarised in the latest edition of Thomson's book (1951, pp.121-123) and have proved of great value to investigators in recent years. They are

i. to use the method of Guilford and Lacey (1947) and stop factorising when the product (here called i) of the two highest factor-loadings of the last factor extracted falls below $1/\sqrt{N}$, i.e. the standard error of zero r.

ii. to use Mosier's (1939) criterion. This criterion is based on common sense: it tells the investigator to keep on factorizing until the algebraic sum of the residuals (excluding the diagonal cells) after as many as possible of their signs have been made positive tends to flatten out. All the factors, before flattening occurs, may be taken as significant

iii. to use Burt's empirical formula for the standard error of each factor loading, viz.

$$(1 - r^2)\sqrt{n/N(n - s - 1)}$$

where r stands for loading, N for number of persons, n for number of tests and s for the ordinal number of the factor. For a factor to be considered significant Vernon advises that at least half its loadings should exceed twice their standard error.

iv. finally, if the results of these three criteria do not agree Vernon recommends the use of McNemar's criterion. This criterion, which Burt (1952) reports arriving at independently, is based on the early claim that an analysis should be continued until the standard deviation of the residuals falls below the standard error of zero r. In developing his criterion McNemar makes use of the fact that factorial residuals when corrected for 'specificity' correspond to partial correlation coefficients. Now, the standard error of a partial r is known, hence the significance of the deviation of factorial residuals from zero could readily be determined. McNemar, however, does not demand that each residual should be corrected for specificity but supplies an approximate procedure. The standard deviation (s) of the residuals is first found. This is then corrected approximately for 'specificity' by dividing it by one minus the average communality, that is by $(1 - M_{h^2})$, of the tests concerned. When this corrected standard deviation (s') reaches or falls below $1/\sqrt{N}$ - the standard error of a zero correlation - no more factors need be extracted. The formula is

$$s' = s/(1 - M_{h^2}).$$

It is clear that McNemar's test is not a test in the usual statistical sense where the probability of a difference having arisen by chance can be stated. With it one simply compares the arithmetic value of an approximated standard deviation with that given by the formula $1/\sqrt{N}$. The lack of statistical sophistication of this procedure need not be emphasised.

More recently Burt (1952), in an encyclopaedic article entitled "Tests of significance in factor analysis", surveys the growth and present state of significance tests. He discusses the empirical and logical rules recommended by psychologists as well as the more rigorous work of Bartlett and Lawley, criticising the latter in rather naive terms. For simple summation and group factor analyses he admits that no rigorous procedures are available, but recommends in particular a chi-square test of his own.

In it chi-square is taken as -

$$(N - 3)\sum(z - \overline{z})^2,$$

with $\frac{1}{2}(n - k)^2 - \frac{1}{2}(n + k)$, degrees of freedom, where n stands for number of tests and k for the number of factors already extracted. Here z refers to Fisher's z-transformation of a correlation coefficient and the $\overline{z}$ is a similar transformation of the amount of the correlation accounted for by the factors. The formula is a generalisation of that used for testing the difference between two correlation coefficients. Justifying it he (Burt, 1952) states that "since the distribution of $z/\sqrt{N - 3}$ is nearly normal, with unit deviation, it follows that on summing the squares for all the discrepancies and multiplying the result by (n - 3), we obtain a quantity which will be distributed approximately as chi-square.

On first appearances Burt's test might appear valid, but as Lawley has pointed out to me in conversation, at least two objections to it can be made either of which would tend to invalidate it. In the first place it is likely that z and $\bar{z}$ would be correlated, in which case a quadratic form of these expressions would be required. In the second place nothing is known about the distribution of the 'correlations' generated by the factor estimates, that is about the part of the formula which after the z-transformation yields $\bar{z}$, and so it is not known whether or not the transformation is appropriate or necessary.

The Significance of a Correlation Matrix.

Our discussion of significance may with advantage begin by referring to a test of the significance of a correlation matrix. One of the commonest criticisms of factor analysts, especially those working in the personality field, is that the size of the correlations they report often tends to be so small, and the reliability of their variables tends to be so low, that it is doubtful whether any valid results can be obtained from a factor analysis of their data. This criticism is sometimes justified. A good demonstration of low correlations is afforded by a study of Bannatyne (1953). The first correlation matrix which Bannatyne considered was obtained with a sample of 95 patients using 14 variables. The variables consisted of scores from the Rorshach test, the Myokinetic test, a word association test, and other projective devices. Only six of the 91 correlations proved to be significantly different from zero, the vast majority of the others being less than 0.1. This situation caused the investigator some concern as he had postulated that two significant factors 'neuroticism' and 'extraversion-introversion' would be found. It was therefore decided to test the correlation matrix as a whole for significance. For this purpose the relevant test (Bartlett, 1950), applicable even for moderately large samples, is -

$$\chi^2 = -(N - \frac{2p + 5}{6})\log_e |R|,$$

where N is the number in the sample, p is the number of tests, and R is the correlation matrix. The degrees of freedom are $\frac{1}{2}p(p - 1)$. To avoid the laborious calculation of $|R|$, resort was had to Lawley's approximate formula, namely:

$$\chi^2 = N\sum r_{ij}^2,$$

or, using Bartlett's multiplier (though the latter is a bit dubious):

$$\chi^2 = -(N - \frac{2p + 5}{6})\sum r_{ij}^2, \quad i > j.$$

The value of chi-square obtained was not significant and Bannatyne was prevented from entering on a fruitless search for the postulated factors.

Assuming then that a given correlation matrix is significant and that one or more centroid factors have been extracted what empirical tests of the significance of the residual matrices, or of the factors obtained, can be relied upon to give reasonably reliable results?

Two tests of the significance of residual matrices, namely Burt's chi-square test and McNemar's test, and two tests of the significance of factors, namely Guilford and Lacey's test and Burt's standard error test, were now checked against the results obtained in the last chapter by Lawley's chi-square test.

Mr. James's Sample ($N = 810$)

Tests of the residual matrix after fitting four maximum likelihood factors gave the following results:-

Significance Test.	Details.
Lawley	$\chi^2 = 129.4$ with 24 d.f. - highly significant
Burt	$\chi^2 = 94.5$ with 24 d.f. - hs " "
McNemar	$s' = 0.076 > 1/\sqrt{N} = 0.035$ - significant.

In agreement with these results Guilford and Lacey's test ($i = .070 > .035$), and Burt's S.E. test showed the fifth factor loadings to be significant.

Tests of the residual matrix in Mr. James's sample after five maximum likelihood factors had been fitted gave the following results:

Significance Test.		Details.
Lawley	$\chi^2 = 63.31$ with 16 d.f.	- highly significant
Burt	$\chi^2 = 21.36$ with 16 d.f.	- N.S.
McNemar	$s' = 0.037 > 0.035$	- just significant.

In Burt's S.E. test a factor was taken as significant if half its loadings exceeded twice their standard errors - following Vernon (1947). But it appears to the writer that this is tooostrict a ruling and he would consider a factor significant if it had two loadings highly significant, provided the tests concerned were not too alike (in the sense of being parallel forms of the same test.)

Mr. Emmett's Sample ($N = 211$).

Emmett in his article (1949) found the residuals after fitting two maximum likelihood factors to be significant at the 0.05 level. Our additional information is:-

Significance Test.		Details.
Burt	$\chi^2 = 1.853$ with 12 d.f.	- N.S.
McNemar	$s' = 0.079 > 1/\sqrt{N} = 0.069$	- significant.

Guilford and Lacey's test when applied to the third factor loadings showed the factor to be significant ($i = .156 > .069$). Burt's S.E. test gave critical ratios for tests 3 and 8 of 2.703 and 5.305 respectively. The corresponding ratios quoted by Emmett, using standard error formulae supplied by Lawley (1949), are 2.35 and 4.79. In agreement with Emmett this factor is taken as significant on Burt's test.

Four factors were extracted from Emmett's matrix by the principal component method and as shown in the last chapter all four were significant. This is one more significant factor than Lawley's test or McNemar's test indicates, and two more than Burt's chi-square test indicates.

Replicated Samples A and B.

Tests of the residuals after fitting two maximum likelihood factors gave the following results:-

Significance Test.	d.f.	Details. 1950A		1951A	
Lawley	13	$\chi^2 = 24.2$	- 5%	$\chi^2 = 20.3$	- N.S.
Burt	13	$\chi^2 = 10.4$	- N.S.	$\chi^2 = 8.5$	- N.S.
McNemar	-	$s' = .101 > .098$	- sig.	$s' = .081 < .082$	- N.S.
		1950B		1951B	
Lawley	26	$\chi^2 = 43.4$	- 2%	$\chi^2 = 50.9$	- 1%
Burt	26	$\chi^2 = 23.8$	- N.S.	$\chi^2 = 23.5$	- N.S.
McNemar	-	$s' = .107 > .098$	- sig.	$s' = .096 > .082$	- sig.

Up to now our comparison of empirical tests of significance with Lawley's test have been carried out on the results of maximum likelihood analyses where efficient estimates of the loadings were available. In practice, however, these empirical tests are generally employed where less efficient estimates got by a centroid or other approximate method has been used. In the last chapter the results of a centroid analysis on James's matrix are reported. When doing the analysis the 'highest r' was always used as 'communality' and no iterations were performed. The results obtained in this way may be considered as approximate only, and it is of interest to apply Lawley's approximate chi-square test and our empirical tests of significance to them and compare the answers with what we know from the maximum likelihood analysis.

Mr. James's Sample (N = 810)

After four centroid factors had been extracted the residual matrix, when tested for significance, gave the following results:-

Significance	Test.	Details.
Lawley	$\chi^2 = 238.2$ with 24 d.f.	- highly sig.
Burt	$\chi^2 = 83.0$ with 24 d.f.	- highly sig.
McNemar	$s' = .061 > 1/\sqrt{N} = .035$	- sig.

In support of these results Guilford and Lacey's test ($i = .070 > .035$) and Burt's S.E. test showed the fifth factor to be significant.

After five centroids had been extracted the results were:-

Significance Test.		Details.
Lawley	χ^2 = 168.9 with 16 d.f.	- highly sig.
Burt	χ^2 = 42.8 with 16 d.f.	- highly sig.
McNemar	s' = .049 > .035	- sig.

The above results compare very favourably with those found on pages 73 and 74. The only discrepancy is that between the results of Burt's chi-square test which shows the centroid residuals after five factors to be highly significant, but the maximum likelihood residuals not to be significant.

Replicated samples 1950A and 1951A

After two centroid factors the results were -

	1950A			1951A		
Test.		d.f.			d.f.	
Lawley	χ^2 = 28.85	13	1%	χ^2 = 27.99	13	1%
Burt	χ^2 = 10.24	13	N.S.	χ^2 = 10.36	13	N.S.
McNemar	s' = .098 < .099		N.S.	s' = .098 > .082		sig.

The chi-square values obtained by the Lawley test here are larger than the corresponding values, namely 24.2 and 20.3, found after two maximum likelihood factors had been fitted which testifies to the superiority of the latter fit.

Mr. Nigniewitsky's Sample (N = 380)

Finally, we have Nigniewitsky's matrix which was factorized by the centroid method only. Here one iteration gave good convergence of the loadings.

The residuals when tested gave the following results:-

Significance Test.		Details.
Lawley	$\chi^2 = 91.0$ with 12 d.f.	- highly sig.
Burt	$\chi^2 = 17.7$ with 12 d.f.	- N.S.
McNemar	$s' = .016 > .051$	- N.S.

Discussion and Conclusions.

In this chapter some empirical rules for telling the analyst when to stop factoring have been described and criticised. Then they have been applied to the residual matrices of our numerical examples after both maximum likelihood and centroid analyses had been performed. Lawley's approximate chi-square test was also applied to the same sets of residuals and the results have been tabulated side by side to facilitate comparison.

In all cases the superiority of the factor fit achieved by the maximum likelihood method is evident, the value of chi-square obtained from residuals using it being smaller than those which resulted when the centroid method was used.

When evaluating the empirical tests of significance interpretation of our findings needs to be done with care. A superficial glance at the results might lead one to conclude that agreement between these and the statistically valid test is fairly good since, in the examples given, they lead to similar decisions in about half the cases. However, a closer look shows that the discrepancies are great. In particular Burt's test grossly underestimates the value of chi-square. To make the point clear the values given by the respective tests are listed below in the order in which they appear in the chapter. When McNemar's test agrees with Burt's a B is entered in the last column, when it agrees with Lawley's test an L is entered; in other cases a sufficient number of factors has not been extracted to decide the issue.

Values of Chi-square.

d.f.	Lawley's test.	Burt's test.	McNemar's test.
24	129.4	94.5	-
16	63.3	21.4	L
13	24.2	10.4	L
13	20.3	8.5	-
26	43.4	23.8	L
26	50.9	23.5	L
24	238.3	83.0	-
16	168.9	42.8	-
13	28.9	10.2	B
13	28.0	10.4	L
12	91.0	17.7	B

This table shows the results given by Burt's test to be in general of a different order from those given by Lawley's. McNemar's test shows up in a slightly better light but other objections to it have already been noted.

In view of these findings neither Burt's nor McNemar's test can be recommended. A much safer procedure, in view of Lawley's recent findings (1955) about the relatively high efficiency of the centroid loadings, would be to apply his chi-square test when using the centroid method. This would necessitate little, if any, more computation than the empirical tests demand.

In general Lawley's test will overestimate the value of chi-square (as a result of the centroid loadings being inefficient estimates) but the fact that this is known is in itself a safeguard. In cases, however, where iteration has been resorted to, as in our 1950A and 1951A samples, the overestimation may not be very large. Our findings after extracting two factors were:

Type of Analysis.	d.f.	1950A Sample. χ^2	1951A Sample. χ^2
Maximum Likelihood	13	24.2	20.3
Centroid	13	28.9	28.0

Chapter V

INTERPRETATION OF FACTORS.

It was noted earlier that the maximum likelihood method provides us with a valid test of the number of uncorrelated factors necessary and sufficient to account for the correlations or covariances between the tests in a battery. Subsequently it was shown that in experienced hands the more widely used but approximate centroid method, with iteration and supplemented by Lawley's chi-square test of significance, could be relied upon to do the job sufficiently well for most practical purposes. Unfortunately, though the dimensionality of the common factor space in any particular study can be decided unequivocally the directions of the axes in it are not, ipso facto, uniquely defined (except in the unidimensional case). Within the factor space the axes may be rotated to any desired position and may or may not be kept orthogonal. The latitude thus allowed has led to considerable difference of opinion regarding the 'correct' approach to the interpretation of factors and many rules have been laid down and semi-analytic methods proposed for attaining greater objectivity. An effort will be made in this chapter to review some of these proposals.

The Aims of Rotation.

Since the introduction of multiple factor analysis the desire to eliminate negative loadings, and achieve parsimony of description by maximizing the number of zero loadings, have been the main directives behind rotation. But parsimony must not be achieved at the expense of clarity and, from a purely descriptive point of view, it is reasonable to expect the position in which the factors are eventually anchored to reflect clearly the structure of the correlation matrix by demonstrating the presence or otherwise of general or bi-polar factors and by throwing into light any clusters or sub-groups of highly correlating tests within it.

A factor pattern, of course, is not independent of the population tested and often the investigator feels he should make the purely descriptive aim subservient to his prior knowledge of the factor composition of the tests if it is desired to maintain some measure of factor invariance from one study to another. All these points call for attention and will be discussed presently; but, before going on, it is right to mention an approach to factor analysis, and the rotation problem known as 'criterion analysis', introduced and advocated by Eysenck, (1950), which is novel and may

The Aims of Rotation.

Since the introduction of multiple factor analysis the desire to eliminate negative loadings, and achieve parsimony of description by maximizing the number of zero loadings, have been the main directives behind rotation. But parsimony must not be achieved at the expense of clarity and, from a purely descriptive point of view, it is reasonable to expect the position in which the factors are eventually anchored to reflect clearly the structure of the correlation matrix by demonstrating the presence or otherwise of general or bi-polar factors and by throwing into light any clusters or sub-groups of highly correlating tests within it.

A factor pattern, of course, is not independent of the population tested and often the investigator feels he should make the purely descriptive aim subservient to his prior knowledge of the factor composition of the tests if it is desired to maintain some measure of factor invariance from one study to another. All these points call for attention and will be discussed presently; but, before going on, it is right to mention an approach to factor analysis, and the rotation problem known as 'criterion analysis', introduced and advocated by Eysenck, (1950), which is novel and may well be worthy of more attention on the part of psychological investigators. Unfortunately, nothing is known as yet about its statistical 'efficiency'.

To introduce the method some digression is necessary. Suppose we wish to demonstrate the existence of some factor like 'neuroticism' in a normal population, a number of tests - measuring what are considered to be neurotic symptoms - which differentiate significantly between a sample of normal people and a sample of neurotic patients are sought. Pure tests of any kind are difficult to come by so the possibility that the tests measure other factors as well as that in which we are especially interested is not ruled out. These tests are then administered to the sample of normal people, and their intercorrelations factor-analysed. The battery is also administered to the sample of neurotic patients and a biserial correlation between the normals and the neurotics found for each test. The latter correlations form a 'criterion' column which is entered alongside the matrix of factor loadings. In this column the highest entries will stand opposite the tests which differentiate best between the groups - the lowest opposite the tests which differentiate least well. It is reasonable then to expect that a factor of neuroticism should have a pattern of loadings the magnitude of which are proportional to the entries in the criterion column. Consequently, a weighted sum of the factors is found which has the property that the sum of the squares of the discrepancies between the entries in it and the corresponding entries in the criterion column is a minimum. From the vector of weights a rotation matrix for rotating the factor loadings to obtain a unique position for the factor of 'neuroticism' can be determined.

Using 'a priori' information.

Our general discussion of the interpretation of factors may conveniently be opened by reference to Emmett's example and his article (1949). A look at his matrix shows that all the correlations are positive. With 209 degrees of freedom a correlation of roughly 0.14 is significant at the 0.05 level so they are all highly significant also. In view of this and the fact that his variables are all measures of cognitive ability it was natural - and, it may be added, in keeping with the traditions of British factor analysts - to postulate a general factor of cognitive ability. From his a priori knowledge of the variables in question Emmett chose variable 3, a non-verbal test, as his best single measure of 'g', and as seen from his rotated factor loadings, Table XVIa, all the common variance of this variable was assigned to the general factor. In this way the positions of all three factors in the factor space were fixed. Factor II proved to be a very clearly defined verbal factor reflecting the cluster of high intercorrelations between the verbal tests, while factor III gave high loadings on the variables which involve 'spacial' ability.

The economy of description which a factor analysis of a correlation matrix affords is well illustrated by this example of Emmett's. In the first place the 9 x 9 matrix has been replaced by a 3 x 9 matrix shown to be necessary and sufficient to account for the co-variation between the variables. The 3 x 9 matrix is then simplified further by rotation so that negative signs are eliminated, many of the loadings are reduced to zero and the factors which result <u>a</u> reflect the 'structure' of the original matrix, and <u>b</u> agree with our <u>a priori</u> knowledge of the psychological nature of the variables concerned.

Perhaps the only valid criticism of the study would be the choice of variable 3 as pivot for the rotation procedure. A different investigator, depending on his idea of the nature of 'g', might wish to choose a different test, or set of tests, to define it and so would arrive at a somewhat different result.

It is hard to see how this subjective element in factor analysis is to be overcome. Clearly the matter is not one for the statistician alone to solve, but it will be well to look, first at the degree of success which analytic or semi-analytic approaches to the rotation problem attain, and later to mention some relevant, but as yet, not well known work by Lawley (195~~6~~7) and Howe (1955) where sampling issues are adequately dealt with.

Thurstone's 'Simple Structure'

One of the most widely used methods for obtaining a psychologically meaningful solution of centroid factors is that proposed by Thurstone as early as 1935. It involves a rotation of the factors into what he calls 'simple structure'. In this position they need be orthogonal no longer and nowadays seldom are. The angles between them, of course, can be calculated.

Before elaborating on the arithmetic, some discussion of the 'simple structure' concept from the psychological viewpoint may be valuable: much of the controversy about psychological factors, and indeed about factor analysis in general, has centered round it. Some of Thurstone's most inspired ideas about mental functioning have - as Professor

Drever has reminded me - got passed over in the rather futile controversies about arithmetical procedure which came to be associated with them.

In the introduction to his book Thurstone (1947, p.58) writes - "our work in the factorial study of the human mind rests on the assumption that mind represents a dynamical system which can eventually be understood in terms of a finite number of parameters. We have assumed, further, that all these parameters, or groups of parameters, are not involved in the individual differences of every kind of mental task. Just as we take it for granted that the individual differences in visual acuity are not involved in pitch discrimination, so we assume that in intellectual tasks some mental or cortical functions are not involved in every task. This is the principle of 'simple structure' or 'simple configuration' in the underlying order for any given set of attributes". Later (p.59) he says of the parameters, or factors, he has in mind, that some of them "may be found to be anatomically determined; others will be physiological; while others will be defined, at first, in experimental, educational, and social terms." In the belief that human behaviour could be adequately described in terms of a limited number of relatively independent concepts, or factors, he set himself the task of discovering what some of these factors were and the fruits of his labours are reported in the psychometric monographs, "Primary Mental Abilities" (1943), "A Factorial Study of Perception" (1944), and numerous other publications. But it is important to note that the discovery of these primary factors was but the <u>first</u> stage in the search he had in mind. The next stage was to try to explain the primary factors themselves. "It is our belief", he writes (1943, p.vi), "that the appraisal of the cognitive and conative primary traits will eventually be accomplished in terms of discriminatory and other rather simple perceptual tasks with individuals in the laboratory and that the tests will look entirely different, superficially, from the traits that they may be found to signify. It may even happen that such abilities as Number, Induction, and Memory may be appraised by tachistoscopically presented discriminatory tasks that do not contain any numbers, that do not call for memorizing in the usual sense, and which do not involve inductive or deductive thinking in explicit verbalized form". A similar stand regarding the value of factors in psychological research has been taken by Eysenck (1955, p.31 et seq.) In his view, "it has always seemed necessary in the investigation of personality to proceed in two stages. In the first place, what seemed to be required was an objectively established dimensional framework which would accommodate the main facts and features of behaviour relevant to mental abnormality." ..... "However, it was also appreciated that this type of static, taxonomic, nosological, dimensional, or classificatory approach would require to be supplemented in due course by a more dynamic or causal type of investigation. If, as the author has argued, statistical factors can be regarded under certain circumstances as causal agents, then it is incumbent upon the investigator not to rest content with the extraction of statistical factors, but to go on to find the psychological or physiological causes indicated by the factor analysis".

The point of view of the factorists is seen in reverse, but not in a conflicting sense, in a paper entitled "Early learning and the percention of space", by Drever (1955). He is investigating Hebb's distinction between early and late learning in the light of some results from an experiment with a sighted and b early and late blind subjects, and says that "certain perceptual abilities having to do with objects in space seem to require a long apprenticeship either in the visual or in the tactile-kinesthetic modalities, and that once the apprenticeship has been served different amounts of later practice have no appreciable effect. "We have in fact, he continues, "something rather like the kind of abilities identified by factorial studies of test performance."

If more evidence of this kind is forthcoming - and here we may recall earlier remarks (Chapter II) concerning localization of function - it will do much to vindicate the factor analyst's position.

With our thoughts orientated by the foregoing remarks let us return to our discussion of 'simple structure'. Thurstone maintained that the identification of factors cannot in the first place be accepted unless they make psychological sense. This implies, amongst other things, that the "factorial description of a test remains invariant when it is moved from one test battery to another". For this reason he was forced to discard the principal component description (Thurstone, 1943, p.91) - "even though the solution it gives is mathematically unique" - in favour of rotation - rotation, guided by a necessity to preserve intact the factorial composition of his tests. The circularity of this argument is obvious, but in a programme of research using antecedent knowledge of the tests employed it was possible alternately to purify the tests and adjust the factors until the goal was in sight.

To understand Thurstone's 'simple structure' it is necessary then to consider the variables employed in one of his studies. Very briefly he aims, when assembling a battery, at excluding tests of a general nature which might have loadings on all his factors. (This implies the avoidance of a general factor, and it is claimed - with some justification - that he was aided in this respect by the populations he employed. These were very often adult populations in which more specialised abilities had had time to develop.) Then, knowing intuitively where his zero loadings should occur, he proceeds mathematically and rotates his factors, guided by graphical considerations, until zeros are obtained in the expected positions.

An Analytical Method for Achieving Simple Structure.

Thurstone's earlier procedure for achieving simple structure by graphical means had a big subjective element and did not always yield a unique solution. More recently he has returned to the problem and has presented an 'analytical method for simple structure' (1953). It is to the latter, and not to the earlier graphical techniques, that attention is given here.

A precise mathematical definition of simple structure, as far as the writer is aware, has never been formulated. The nearest approach to it is that given by Tucker (1955) but it is beset about by so many restrictions that it is almost impossible to imagine data to which it would be applicable.

In the article by Thurstone (1954), referred to above, the approach is one of defining hyperplanes one at a time. If the tests have been properly chosen there should be as many distinct hyperplanes as the number of factors extracted. To define a hyperplane some test is chosen which has many low correlations with other tests in the battery but a few high ones. The direction cosines of this test give the first approximations to the direction cosines to the normal to the hyperplane in question. These approximations are improved by a process of successive application of predetermined weights to the projections of the other tests on that selected until eventually a stable position for the normal to the hyperplane is found.

This method of Thurstone's, although not well suited to a matrix which has a clearly defined general factor was applied to Emmett's loadings. The results are given in table XVIc, the angles between the rotated factors being:

	II	III
I	120.0°	96.0°
II		98.4°

Later in this chapter Thurstone's analytical method is applied to two studies better suited to its use, but even here it must be claimed that it is not altogether without merit. Factors I and II become oblique with a correlation of -0.501. The latter is a well defined 'verbal' factor. Factor III, which is virtually uncorrelated with factors I and II, looks like a 'space' factor with loadings agreeing in some measure with those found by Emmett. Factor I, formerly labelled 'g', has suffered from the rotation and is now rather difficult to define. Since it has been severely purged of its verbal content it may perhaps be labelled 'reasoning'.

The Quadrimax Method.

Another analytical method which purports to objectify the rotation procedure is the 'quadrimax method', (Neuhaus and Wrigley, 1954). It is mentioned because it has some support amongst psychologists, but Dr. Lawley admits being shocked by it.

Regarding its development the authors say that "it provides a rather interesting case of four or five investigators working independently, in ignorance of the work of the others, and reaching the same results". (The other writers referred to are Carroll (1953), Ferguson (1953) and Saunders (1953)).

A detailed account of the method would be too tedious to include here but in general it aims at arriving at a parsimonious description of the factor results. In the writers' own words it "starts with any factoring of the correlation matrix and attempts, using orthogonal transformations, to increase those loadings which are already large and to reduce those which are already small to as near zero as possible. In other words, the aim is to increase inequalities between factor loadings as much as possible and still keep a matrix which will reconstitute the correlations". It is hard to see how such an arbitrary procedure could seem justified to psychologists looking for 'meaning'.

In practice, application of the method requires taking the factors two at a time, and by a straightforward though lengthy calculation, finding the angle through which they must be rotated in order to maximise the fourth powers of the loadings. If there are k factors then there are in the first instance $\frac{1}{2}k(k - 1)$ pairs of factors to be considered and this is said to constitute a 'cycle'. When one 'cycle' of rotations has been run through a second 'cycle' is commenced and so on until the sum of the fourth powers of the loadings converges.

With a large number of factors the method is laborious without the aid of an electronic computer, but convergence is rapid and for only a few factors the work can be carried out with a desk computer. The method allows for a general factor if one exists. It was applied to the maximum likelihood loadings in Emmett's example and the results are given in table XVIb.

The first rotation matrix which the method yielded was

$$\begin{array}{rr} .716 & .698 \\ -.698 & .716 \end{array}$$

for factors I_0 and II_0. After this rotation had been performed factors I_1 and III_0 were considered, but the angle for the rotation was found to be $89^\circ\ 7.5'$, i.e. almost 90°, so the rotation was not performed. Finally factors II_1 and III_0 were considered and the angle for this rotation was found to be $87^\circ\ 56'$. It was decided that this rotation also would make no noticable difference to the loadings so it was not performed.

A comparison of the quadrimax rotated loadings table XVIb, and Emmett's rotated loadings table XVIa shows the superiority of the latter interpretation both as regards the signs and magnitudes of the loadings. Nevertheless, the quadrimax method does succeed in bringing out the salient points in the analysis. Neuhaus and Wrigley too admit that it may often be beneficial to supplement this method by the well known graphical methods once the general overall picture has been revealed.

The Principal Components of Emmett's Matrix.

No attempt was made to rotate the four principal components found for the Emmett matrix. As is well known all the principal components of a matrix are required to reproduce the correlations. Yet it is sometimes thought that since the first few principal components generally account for a great part of the variance of the tests - in our 9 x 9 matrix the first four account for 81.5% of the total variance - later components may be disregarded in the search for psychological meaning. That this can be a very dangerous procedure is demonstrated with Emmett's matrix. In table VIIa the column sums of the correlation matrix - excluding diagonal cells - are given. These give a rough measure of the extent to which any test in the battery correlates with the remaining tests. It would be natural then to expect the column sums to correlate positively and to a high degree with the test communalities which result from the factor analysis. In the Emmett example the rank correlation between the column sums and the communalities after three maximum likelihood factors is 0.63, but the rank correlation between the column sums and the 'communalities' after four principal components is -0.52. One can only conclude that the four principal components give a very inadequate reproduction of the correlations and that the use of unities in the diagonal cells of the matrix is the main cause of the distorted picture. The principal component method, of course, can be employed with communalities. This is equivalent to Burt's 'weighed summation method' with reduced self-correlations in the diagonal cells and the method is known to give very satisfactory results.

The 1950C and 1951C Matrices.

The quadrimax method and Thurstone's analytical method for 'simple structure' was now applied to the centroid loadings for the replicated 1950C and 1951C matrices. To make a comparison of the results easier centroids II and III in the 1951 matrix were interchanged so as to correspond with the 1950 centroids.

The results of the 1950 centroids after rotation are shown in table XVIIa and of the 1951 centroids in table XVIIb. It should be noted that the quadrimax rotations retain the orthogonality between the factors, but Thurstone's method allows the factors to become oblique. The angles between the respective factors in the latter case are

	1950C		1951C	
	II	III	II	III
I	96°24'	88°36'	97°12'	73°42'
II		115° 6'		109°24'

The time taken to carry out the rotation by either the quadrimax or Thurstone's method was approximately 5 hours using a desk computer.

In the table below the order in which the rotations in the quadrimax method were performed, the angles of rotation and the sums of the squares of the fourth powers of the loadings are given. Rotations were stopped as soon as it became apparent that no further noticable advantage would be achieved.

1950C

Factors rotated	I_0-II_0	I_1-III_0	II_1-III_1	I_2-II_2
Angle	50°34'	80°59'	54°59'	1°56'
Sum of (loadings)^4	1.7994	1.8297	2.2235	-

1951C

Factors rotated	I_0-III_0	I_1-II_0	II_1-III_1	I_2-III_2	I_3-II_2
Angle	48°32'	43°22'	14°2'	53°23'	89°38'
Sum of $(\text{loadings})^4$	1.6854	2.5043	2.6422	2.7972	-

Though complete convergence of the sums of the fourth powers of the loadings has not been reached by the rotations performed the tendency towards convergence is clear.

An attempt must now be made to evaluate the results given in table XVII. Even though the Thurstone factors are slightly oblique the results obtained by the two methods of rotation bear a clear-cut resemblance. Both methods have succeeded in revealing the 'structure' of the respective matrices. In each case the first factor has high positive loadings on the cognitive tests, while the third factor - defined by tests 11, 14 and 15 - is probably a 'speed' factor.

When comparing results from replicated experiments the fact that the quadrimax method allows for a general factor while Thurstone's method is more appropriate in its absence should be remembered. But having chosen a suitable method one would expect to find fairly close resemblance between the factor structures in such experiments. Nevertheless, as Bartlett reminds us (1953, pp.76-77), "it is too much to expect complete invariance from population to population. A variation would not invalidate the idea of a true underlying structure - we have the example of genetic characters, which occur in differing proportions in different populations, and yet the genetic structure is established. The statistician will not thus expect complete invariance".

If a reasonable attitude like this is adopted it can be claimed that the factors which have evolved from the 1950C and 1951C matrices given in table XVII correspond.

As a further check on this correspondence the rotated loadings for the 1950C and 1951C analyses, found by the Thurstone method, were compared by Ahmavaara's (1954) transformation procedure (reviewed by Maxwell, 1955). This involves rotating one set of loadings into maximum conformity, in the least square sense, with the other. The comparison matrix found by this procedure is -

		1951C		
	FACTORS	I	II	III
	I	.844	.343	-.161
1950C	II	-.112	.929	.029
	III	-.426	.102	.689

where the entries in the diagonal cells show the correlations between respective factors. The correlations - especially those between factors I and II - in each study are very high and give further support to the claim that there is a good measure of 'factor invariance' in this replicated study. But invariance is the special topic dealt with in the next chapter so nothing more need be said about it here.

Classifying Tests by Successive Bifurcation.

So far in this chapter two semi-analytical methods which aim at objectifying the rotation procedure in factor analysis have been considered. They are empirical methods only and their authors do not make any exaggerated claims on their behalf. But the illustrations of their use which have been presented both here and by their respective authors show that the methods have some merit. In each case they yield results which when referred back to the original correlation matrix do reflect its structure. It must be remembered that the correlation matrices presented here - for reasons of space - are much smaller than matrices generally met in practice. It is in this light that such analytical methods must be assessed, for every experienced factor analyst knows how difficult it is to get even a crude idea of the structure of a large correlation matrix however assiduously it is inspected with the naked eye.

But once its basic structure is revealed the experimenter is in a much better position to classify his tests, describe his results and check on the degree to which his expectations or hypotheses have been confirmed.

Other methods of classifying tests in the light of their intercorrelations have been referred to in an earlier chapter (Thomson 1939, 5th. ed., p.20 et seq.; Burt, 1950, pp44 &45). The method which Burt advocates is to take the unrotated factors in turn and dichotomise the variables according to the signs of the loadings. The writer is not convinced that this is the best method to use on all occasions. What in effect it does is isolate variables which lie in the same quadrant, octant or hyperquadrant; but this does not necessarily isolate 'clusters', for a centroid or a principal component tends to pass through rather than flank a cluster of tests which are highly intercorrelated.

Burt's Group Factor Method

At this stage it is appropriate to say something about Burt's group factor method. Having classified his tests by a process of dichotomisation using the results of a simple summation (or preferably, he would claim, a weighted summation analysis) Burt rewrites his correlation matrix in the order in which the tests have been classified by his bifurcation procedure. In this way the structure of the matrix is much easier to assess. If a general factor is present it is extracted first, and later, attention is given to the group factors. Overlapping factors too can be dealt with.

The details of Burt's mathematical procedures, which are approximate only, need not be given here. They can be found in his book (1940) and in various articles (1949, 1950). His methods, however, deserve attention for they have been found in the past to give clear results provided fine precision is not called for. They have one great advantage - that of keeping the experimenter's attention fixed on the correlation matrix which he is trying to describe, whereas with the other techniques mentioned there is a tendency to get lost in a maze of calculations.

Our 1950A and 1951A matrices are well suited to a group factor analysis and Burt's methods have been applied to them. As the matrices stand, tables X and XI, the tests are already classified, the five neurotic variables appearing first followed by the three cognitive variables. From the matrices it would appear that a general factor in each case should be postulated as the two groups of tests are uncorrelated. Over and above this general factor there would appear to be two group factors one for the cognitive and one for the neurotic variables. The results of the group factor analysis are given in table XVIIIa.

For purposes of comparison the quadrimax method, which allows for a general factor, was also applied to the 1950A and 1951A results and the rotated factor loadings are given in table XVIIIb. Here it would seem obvious to conclude that Burt's group factor analyses give a clearer picture of the covariation between the variables than does the quadrimax method. But there is a statistical difficulty. The tests of significance reported in the last chapter showed the residuals after two factors in the 1951A matrix not to be significant and the factor space was unequivocally shown to be of two dimensions. With respect to this matrix then the group factor method has an unfair advantage over the quadrimax method. Given three dimensions within which to move the latter might do as well as the former.

Regarding the 1950A matrix the residuals ~~on~~ for the maximum likelihood and McNemar tests were just significant after two factors (though Burt's own test would not admit the fact). In this case it would be justifiable to consider the factor space as being of three dimensions and the results given in table XVIIIa for the 1950A rotated loadings would be justified.

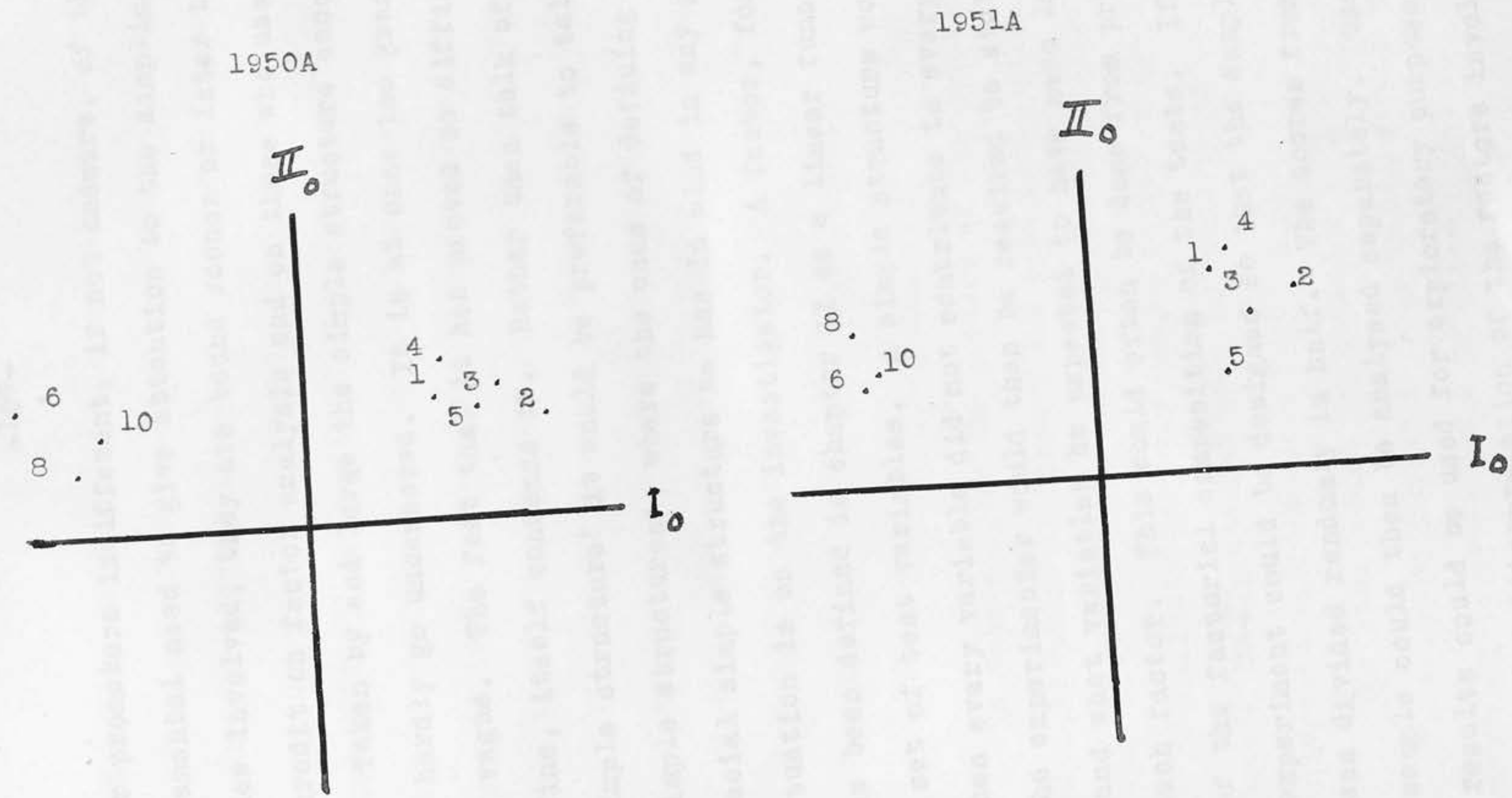
Plots of the factor loadings for the two matrices are given in figure I. From them it is clear that two oblique or correlated factors would describe the results very well in the 1951A case where a third dimension is not significant. The most appropriate analytical technique to apply would then be that of Thurstone, but in this very simple example the results are so obvious that the calculation itself is unnecessary.

The Statistician's Approach to Simple Structure.

The writer, from his practical experience of factorial work, is unwilling to condemn outright the analytical approaches to the interpretation of factors discussed above. When faced with a large number of factors extracted from a large matrix of almost unknown psychological content, as in a pilot study, these approaches can be of value in the initial effort to classify the variables and locate roughly the principal dimensions involved. On the other hand it is realised that from the statistical, if not from the common sense point of view, one must be critical of them. In as far as the methods are empirical only and

FIGURE I

Plots of Factors



their promotors indifferent, if not unaware, of the fundamental need to give attention to the sampling issues involved, they are bound sooner or later to bring discredit on factor analysis and on those who use it.

Taken by and large the simple structure concept too can hardly go uncensored. It is at once too grandiose and vague. The fact that it has proved so difficult to define, itself condemns it. Rather than talk of 'simple structure' it would be preferable to talk of 'simple structures', where the onus of defining the special simple structure he has in mind in any given situation is on the investigator. A factor, for example, has been defined in chapter II as a linear function of a set of test variables. A simple structure would exist when every variable did not contribute to every factor. The experimenter would then be required to state beforehand what variables he expected to have zero loadings on each factor. This could often be done from prior knowledge of the factorial composition of the tests. If not the experiment could be designed so that the sample of testees was divided randomly in half. The scores from each sample could then be analysed separately. One set of results could be used for exploratory purposes. A tentative interpretation of the factors involved could be obtained from it by graphical and other means. Once a probable simple structure had thus been outlined it could be tested on the second sample. What makes such a procedure attractive to-day is the fact that recent work by Howe (1955) and Lawley (1956) shows how the test can be made. These writers, working independently, have developed valid large sample statistical techniques for estimating factor loadings and testing their fit to the correlation matrix when the number and location of the zero loadings on each factor is specified beforehand. Details of their work were not available in time to receive full attention in this thesis, but it is wholeheartedly recommended.

Summary.

In this chapter the problem of the psychological interpretation of factors has been dealt with. It is shown that while the subjective element involved is great it is unlikely that any single objective criterion for factor interpretation will be found which will give universal satisfaction.

The aims behind the rotation of factors are discussed and a study by Emmett is used to illustrate the classical approach to the problem of interpretation.

Next the plausibility of the existence of primary factors and the possibility of accounting for their development and function is reviewed. At this stage Thurstone's 'simple structure' idea is briefly described and evaluated. An analytical method for achieving 'simple structure', and the 'quadrimax method' for interpreting factors objectively are then presented and applied to Emmett's example and to the 1950 and 1951 samples. From an examination of the results it is concluded that both methods have some merit for revealing the underlying structure of a matrix and might be of value when dealing with a large one the structure of which is not immediately apparent.

TAKING ~~Using~~ Emmett's matrix as an example a warning is given about the dangers involved in using the first few principal components as a means of interpreting the covariation between a set of variables, attractive though such a procedure might appear in view of the large amount of variance extracted by these components.

Burt's bifurcation method of classifying tests is then critically examined and his method of finding group factors mentioned. The latter is applied to the maximum likelihood factors for the 1950A and 1951A matrices and the results compared with those obtained by applying the quadrimax method to the same results. Here Burt's group factor approach - though his techniques, statistically speaking are approximate only - is shown to be useful. The method too has the advantage of keeping the experimenter's attention concentrated on the correlation matrix which is being described - rather than on the arithmetical procedures involved.

Finally an effort is made to assess present rotation procedures, and analytical techniques. The latter are not wholly condemned, but the complete neglect of sampling issues by their promoters must not be allowed to go uncensored.

Thurstone's 'simple structure' concept too is severely criticised for its psychological vagueness and its statistical imprecision. The idea of 'simple structures' (in the plural) is proposed as an alternative. By these the writer means patterns of factor loadings where the number and location of zero entries are predetermined. If such patterns cannot be made on prior knowledge of the factor composition of the tests, the factorial experiment can be designed so that probable patterns can be a) obtained, and b) tested. Recent large sample statistical methods by Howe and by Lawley for dealing with the problems involved are welcomed.

Chapter VI

COMPARISON OF FACTORS IN REPLICATED EXPERIMENTS.

The problem of deciding whether or not a factor obtained in one investigation may be legitimately identified with a factor obtained in a second and similar investigation, is one of the most complex in factor analysis. The problem has many aspects and many sources of variation are involved.

The most difficult situation presumably is one in which different tests as well as different individuals have been employed in the two studies. Here a comparison of factors would appear reasonable only if the tests defining them in the two studies are known, on a priori grounds, to be alike in psychological 'content'.

A less hazardous situation is one in which some or all of the tests administered to the two samples are the same. Here a major question is that of the homogeneity of the samples, but even if they were found to be unhomogeneous the experimenter would not find himself in completely unknown territory, for the effect of 'selection' on factor loadings has been dealt with by both Thomson (1939) and Thurstone (1947).

Apart from tests and populations there is too the question of whether factor scores or factor loadings should be compared. Happily this question partly answers itself, for it is clear that a comparison of factor scores is possible only when the same people are concerned in the two studies. In this case the correlation between their factor scores on the two occasions - whether or not the two batteries of tests are identical - will give a measure of comparison between corresponding factors. Estimation of factor scores, however, is a laborious business and a method for comparing factors by more direct means would repay research.

The 'Unadjusted' Correlation Between Factor Loadings.

Empirical methods for comparing factors have recently been reviewed by Barlow and Burt (1954) in a note in the British Journal of Statistical Psychology entitled, "The identification of factors from different experiments". They recommend the correlation of factor scores where this is possible, but also lay stress on a 'coefficient of similarity', or 'unadjusted' correlation coefficient between two sets of factor saturations (Burt, 1935, pp. 245 - 314).

The formula for the 'unadjusted' correlation is given as -

$$s_{12} = f'_1 f_2 / \sqrt{f'_1 f_1 \times f'_2 f_2}$$

where f_1 and f_2 are the respective vectors of factor loadings, and f'_1 and f'_2 are the corresponding transposed vectors.

The formula is derived from certain properties of principal component axes. When the loadings obtained from a principal component analysis are normalized by columns the direction cosines of the axes with the reference to the orthogonal test vectors are obtained. Since the components are orthogonal it follows from the cosine rule that the inner product of the loadings of any pair of components is zero. Moreover, if the same tests have been used in two studies and the resulting principal components are referred to the same set of orthogonal test axes then corresponding components will tend to coincide and the inner product of their direction cosines will tend to unity.

It is clear that an 'adjusted' correlation can be obtained only when the same tests have been used in each study.

Barlow and Burt state that a product-moment correlation coefficient between the saturations of two factors is not an appropriate measure of the resemblance between the factors. Such a coefficient, they claim, implicitly compares, "not the saturations but their deviations, thus changing the smaller saturations into negative deviations. And, if the incidence of these negative deviations differs in the two experiments, they may easily produce a negative correlation: they may do so, even when the measurements for the two factors are almost exactly the same."

In the table below the 'unadjusted' correlation coefficient, and the product-moment correlation coefficients between the unrotated maximum likelihood factors for the 1950A and 1951A studies are given. For convenience the 1950 factors are denoted by I and II and the 1951 factors by I' and II' respectively.

	I I'	II II'	I II	I' II'
Unadjusted correlations	.990	.868	.185	.130
Prod.-moment "	.997	.213	.267	.541

The results support Barlow and Burt's claim that the 'unadjusted' correlation coefficient gives the more reasonable results. However, the number of variables in these studies is so small that it would be unwise on so little evidence to come to any definite conclusions about the respective merits of the two coefficients.

Ahmavaara's Transformation Analysis

More recently the problem of "how to make exact comparisons between different factorial studies carried out with different experimental populations", has been attacked by Ahmavaara (1954). Firstly, he demonstrates a striking parallel between Pearson's selection theory and Thurstone's multiple-factor theory. Then by postulating 'a theory of factorial invariance', and using an argument similar to that used by Pearson he constructs a 'comparison matrix' by which the results of two studies can be compared. More precisely suppose there are two n x r matrices, X and Y, of factor loadings one for each of two samples, where n refers to tests and r to factors. Then if the matrix X be taken as the invariant link between the factor scores for the two samples - corresponding to the invariance of the regression plane assumed by Pearson - and the specifics remain orthogonal to the common factor and to each other, the comparison matrix takes the form $(X'X)^{-1}X'Y$.

This transformation matrix is identical with that at which we arrive if we set out to rotate the matrix X into maximum conformity, in the least square sense, with the matrix Y. The problem is one of finding an r x r matrix, which may be called H, such that

$$F = (XH - Y)'(XH - Y) \qquad \text{....(1)}$$

is a minimum. Expanding equation (1), we get -

$$F = (H'X'XH - 2Y'XH + Y'Y) \qquad \text{....(2)}$$

Equation (2) must now be differentiated with respect to H and the result equated to zero.

$$\delta F/\delta H = 2X'XH - 2Y'X,$$

which, when equated to zero, gives

$$X'XH = Y'X$$

or, $$H = (X'X)^{-1}X'Y \qquad \text{....(3)}$$

This is the transformation matrix required.

But it is clear that the transformation could be performed in the opposite direction in which case the matrix Y would be rotated into maximum conformity with the matrix X. The transformation matrix is now given by

$$K = (Y'Y)^{-1}Y'X \qquad \text{....(4)}$$

The values of H, equation (3), and K, equation (4), were calculated for our 1950A and 1951A maximum likelihood factor loadings, the former factor matrix being taken as X and the latter as Y. The resulting values are -

$$H = \begin{bmatrix} .908 & .036 \\ -.137 & 1.274 \end{bmatrix}, \qquad \text{and } K = \begin{bmatrix} 1.080 & -.014 \\ .133 & .715 \end{bmatrix}$$

These matrices, when normalized by columns, give the correlations between the respective factors in the rotated positions. They are -

$$\begin{array}{ccc} & I' & II' \\ I & .999 & .040 \\ II & -.107 & .994 \end{array} \text{, and } \begin{array}{ccc} & I & II \\ I' & 1.000 & -.013 \\ II' & .183 & .983 \end{array}$$

A look at these correlation matrices is sufficient to show the close resemblance between the factor loadings in the two studies. The results too agree well with the 'unadjusted' correlation coefficients between the factors reported earlier in this chapter.

Ahmavaara's method has the considerable advantage that it can be employed when the two studies have only some of their tests in common. It can also be applied to factors after they have been rotated, whether or not they are orthogonal in their rotated position. An example of its application to oblique factors has already been reported (chapter V, p.$\overset{100}{75}$).

A New Approach

The methods so far discussed for comparing factors obtained from different experiments, being empirical only, cannot be considered wholly satisfactory. This fact led the writer to attempt an approach to the problem along more orthodox statistical lines. A simple type of situation only is dealt with, namely, that in which the same battery of tests has been administered to two comparable and independent samples as is the case with our 1950 and 1951 replicated data.

As mentioned earlier one of the big disadvantages of the more commonly used factorial techniques, (and here the principal component method might be included) is that they are not invariant under changes of scale. In short the loadings obtained if the covariance rather than the correlation matrix is analysed are not proportional. Lawley's maximum likelihood method I is the only existing factor method which does not suffer from this defect. Loadings obtained by it using the correlation matrix have only got to be multiplied by the standard deviations of the respective tests to get the loadings which would have resulted had the covariance matrix been employed.

With the centroid, and principal component analyses of correlation matrices the 'standardization' implicit in the correlation coefficients (which in essence are standardized covariances) is automatically and irretrievably carried over when factor scores are obtained. This changing of yard-stick immediately prejudices any attempt to compare factors obtained from different studies, since the 'standardization' has already resulted in equating to unity the variance of each test in each experiment.

Thurstone claims that his 'simple structure' is invariant under change of scale and Ahmavaara (1954, p.32) supports him. But in Thurstone's case there is always the criticism that 'simple structure' is not itself rigorously defined so that the claim does not carry with it the conviction it would otherwise do.

The unsatisfactory nature of the 'standardization' procedure has long been recognised. Thomson discusses it in the final chapter of his book, and Burt (1940, p.264) makes a plea for the use of the variance-covariance rather than the correlation matrix. Thomson (1951, p.330) and Hotelling (1933, p.510) too suggest methods by which a system of rational units for the more commonly used tests might be derived.

At the recent Symposium at Uppsala (1953) the evils of standardization of test and factor scores again came under fire, and general concern was felt about present procedures. It would appear then that until a standard metric for the more commonly used tests is agreed upon it would be well when comparing factors from different experiments to avoid standardization altogether and work with the variance-covariance rather than with the correlation matrices. This is the procedure followed below.

Comparing Two Variance-Covariance Matrices.

The need for a valid statistical test of the difference between two variance-covariance matrices is obvious for if two such matrices do not differ significantly then it must follow that the factors derived from them can be equated.

The test employed here follows from work by S.S. Wilks (1932, pp.471 et. seq.). The writer is indebted to Dr. D.N. Lawley for the following statement of it and notes his warning that it is valid for large samples only. Our 1950 and 1951 samples are only of medium size, 104 and 148 respectively, but they will serve to illustrate the use of the test.

It should be mentioned that the calculations involved are laborious since three large determinants have to be evaluated, but this will not be a major difficulty when electronic computing facilities become more readily available. Using a desk computer the writer found that a 10x10 determinant could be evaluated in approximately seven hours by the pivotal condensation method provided the inverse of the matrix is not also required.

The Test

Let the two samples have degrees of freedom n_1 and n_2 respectively (i.e., the sizes may be taken as (n_1+1), and (n_2+1). Let A_1 and A_2 be the respective covariance matrices with p variates each.

Let the 'pooled' covariance matrix be A, with (n_1+n_2) degrees of freedom: i.e.,

$$A = \frac{n_1A_1 + n_2A_2}{n_1 + n_2}.$$

Then the quantity

$$(n_1+n_2)\ln|A| - n_1\ln|A_1| - n_2\ln|A_2| \qquad \ldots\ldots (5)$$

where $\ln \equiv \log_e$.

is distributed, on the null hypothesis, approximately as chi-square with $\frac{1}{2}p(p+1)$ degrees of freedom where p refers to tests. Considerations by Box (1949) show that this test can be made more sensitive in the case of moderately large samples by applying the multiplier,

$$1 - \frac{2p^2+3p-1}{6(p+1)}\left(\frac{1}{n_1}+\frac{1}{n_2}-\frac{1}{n_1+n_2}\right)$$

to the expression for chi-square given above.

An Example

When evaluating the determinants of our 1950 and 1951 covariance matrices variables 7 and 9 were placed last. In this way it was possible to find the values of the determinants of these matrices with these variables excluded, as in our 1950A and 1951A matrices, and also with them included as in our 1950B and 1951B matrices. It was also possible to test for a significant difference the matrices with the neurotic variables (i.e. variables 1 - 5) only included.

The values of chi-square obtained by the use of the above formula, (using the Box multiplier) together with their respective degrees of freedom and levels of significance, are given below:-

Covariance matrix	chi-sq.	d.f.	Sig. level.
Variables 1 - 5 only	22.3	15	N.S.
Variables 1 - 6, 8 and 10	81.6	36	highly sig.
Variables 1 - 10	107.1	55	highly sig.

Bearing in mind the small size of the samples the results show that whereas the covariance matrices do not differ if only the five 'neurotic' variables are considered, they differ very significantly when the cognitive variables, 6, 8 and 10, and 6, 7, 8, 9, and 10 are included.

These findings are in accord with the results given in chapter IV. There the residual matrices, after two maximum likelihood factors had been fitted were shown to differ - for whereas 1950A residual matrix was significant at the 0.05 level the 1951A residual matrix was not significant, and whereas the 1950B residual matrix was significant at the 0.02 level the 1951B was significant at the 0.01 level (p. 75). However, the high level of significance of the differences between our covariance matrices found above and the not so pronounced differences demonstrated between the residual matrices strongly suggest that differences of medium degree also exist between the factors themselves. Attention will now be concentrated on this latter problem.

Significant Differences Between Factors.

When two covariance matrices differ significantly as, for example the 1950A and 1951A matrices do, it is reasonable to enquire whether this difference is reflected in any, all or only some of the factors derived from them. This leads NOT ONLY to the question of how to test the significance of the difference between factors by an examination of their respective loadings, but also to the question of whether or not this is a ~~reasonable~~ SENSIBLE thing to do.

Factor loadings derived from a centroid or maximum likelihood analysis are not regression weights for finding factor scores. They are not even proportional to such weights, as is the case with principal component loadings, though they are highly correlated with them. Nevertheless, given the variances and covariances of two sets of factor loadings there could be no objection to testing the differences between such loadings for significance by routine statistical methods provided the necessary assumptions were fulfilled.

The loadings could be compared two at a time, but an overall test which would test all the loadings of one factor simultaneously against all the loadings of another and corresponding factor could also be employed. Such a test will now be outlined. Unfortunately it suffers from the disadvantage that it involves the inversion of a pxp matrix, where p is the number of variables concerned.

Significance of the Difference between Factor Loadings, and Factors obtained from Independent Samples

Suppose a battery of p tests is administered to two independent random samples, and the two pxp matrices of intercorrelations of the tests are factor-analysed by the maximum likelihood method. Call the loadings on the first factor fitted in each case $\hat{l}_{ir}$ and $\hat{l}'_{ir}$ respectively where i (1.....p) refers to tests and r refers to factors. In this case r = 1. Now if an estimate "a" of the variance of $\hat{l}_{ir}$, and an estimate "a'" of the variance of $\hat{l}'_{ir}$ are available the difference between $\hat{l}_{ir}$ and $\hat{l}'_{ir}$ can be tested for significance by the formula,

$$\text{C.R.} = (\hat{l}_{ir} - \hat{l}'_{ir})/(a + a')^{1/2} \qquad \ldots.6$$

Here it is assumed that $\hat{l}_{ir}$ and $\hat{l}'_{ir}$, or more strictly the differences between them, are normally or approximately normally distributed.

The argument can now be generalized so that all the loadings on one factor from one sample can be tested simultaneously against their counterparts from the other sample, that is the two factors can be tested against each other.

Let the variance-covariance matrix for the factor loadings from the first sample be represented by the matrix a_{ij} (i,j = 1.....p) and the corresponding matrix for the second sample by a'_{ij}. Then, when the sample sizes are equal, the matrix $(a_{ij} + a'_{ij})$ which may be called b_{ij}, is the 'within groups' variance-covariance matrix. (When the samples are unequal in size a weighted sum of the respective matrices can be used).

The expression -

$$\left\{(n_1 + n_2 - r - 1)/(n_1 + n_2 - 2)\right\}\theta^2$$

is now calculated, where

$$\theta^2 = n_1 n_2/(n_1 + n_2)\sum_i^p \sum_j^p b^{ij} d_i d_j, \qquad (i,j = 1.....p)$$

b^{ij} is the matrix inverse to b_{ij}, $d_i = (\hat{l}_{ir} - \hat{l}'_{ir})$, and similarly for d_j: n_1 and n_2 are the sample sizes. This expression is distributed as chi-square with r degrees of freedom.

The Standard Errors of Factor Loadings.

Various approximate formulae for the standard error of a factor loading have from time to time appeared in the literature (Burt and Banks, 1947; Burt 1952; Holzinger and Harman, 1941, pp. 124-132; Kelley, 1947;). Burt's formula is perhaps best known and has already been discussed in Chapter IV in reference to tests of significance of factors.

A more obvious but also very rough formula for the standard error of a factor loading was suggested to me (strictly sotto voce) by Dr. Lawley. It is $\sqrt{V_s/N}$, where N is the size of the sample and V_s is the residual variance of the test after the extraction of factors

1 to s. The formula could be applied to loadings either before of after rotation. It will be referred to as the S.V. formula. Standard errors calculated by it and by Burt's formula for the second factor from the 1950A and 1951A analyses are given in table XIX. The figures show wide discrepancies, and the values obtained by Burt's formula are consistently bigger than those got by the S.V. formula. The moral to be learnt from the comparison is not to put much confidence in empirical formulae like those in question.

In 1953 Lawley (Uppsala Symposium) tackled the problem of finding large-sampling error formulae for the estimates of factor loadings. The approach assumes that either the values or the ratios of the specific factors are known and Lawley himself (personal communication) considers the approach to be of theoretical interest rather than of practical value. However to see what results got by Lawley's equations might look like the ratios of the 'specifics' obtained in the 1950A and 1951A studies were taken as the true values of these quantities for the respective populations and the necessary calculations performed. The procedure is reported below in some detail as the results obtained seem to the writer to be of interest.

The formula given by Lawley for the covariance of the loadings $\hat{l}_{ir}$ and $\hat{l}_{jr}$, where r refers to factors and i and j to tests, is somewhat unpleasant. It is

$$L_r/n(L_r - \sigma^2)\left[c_{ij} - \frac{L_r}{2(L_r-\sigma^2)}\, l_{ir}l_{jr} + \sum_{m \neq r} \frac{L_m}{L_m-\sigma^2}\left(\frac{(L_r-\sigma^2)2}{(L_r - L_m)^2} - 1\right) l_{im}l_{jm}\right] \quad .....(8)$$

plus an additional term in the case where σ^2 has to be estimated. (In our case this additional term was found to be insignificant so it need not be added here). In this formula c_{ij} stands for C, the population variance-covariance matrix, with n degrees of freedom based on a random sample of (n + 1) individuals. L_r is the r^{th} latent root of C. The ratios of the 'specifics' are assumed known and the units of measurement of the variables are chosen so as to make all the specifics equal to some constant, σ^2.

Using Lawley's Formula.

To apply the formula the following rather circuitous path was taken the only justification for which may well lie in the results to which it led.

1). The variables in the 1950A matrices were considered and variable (8) which had the highest reliability coefficient was chosen. The specific variance of this test in the 1950A analysis was 0.599 (based on 103 degrees of freedom) and in the 1951A analysis was 0.362 (based on 147 degrees of freedom). A weighted sum of these two specifics gave the value 0.460. This is the value of σ^2 used below.

2) Since the loadings in a maximum likelihood analysis (Method I) are invariant under change of scale the scales of factor loadings for the two studies and the correlation matrices from which they were derived were now altered so as to make the specifics for each test in each analysis equal to 0.460. This put the data in a form suitable for the application of equation (8), (if the estimated values are substituted for the population values of the several parameters).

The adjusted factor loadings are given in table XX.

3) Estimates of the latent roots L_r, of the matrices were obtained by the equation -

$$L_r = l_r' l_r + \sigma^2 \quad \ldots\ldots(9)$$

Estimates of all the quantities required by equation (1) were now available so the variance-covariance matrices of the adjusted factor loadings (table XX) were calculated.

These are given for the first factor of each study in table XXI.

At this stage formula (7) could have been employed to compare these factors, but it was decided to be less ambitious and compare corresponding pairs of loadings in the two studies instead. To this end table XXII was drawn up where critical ratios, obtained by equation (6), are given. In table XXIIa corresponding data for the second factor from each study appear.

A glance at tables XXII and XXIIa shows that the two sets of factors are in close agreement. For factor I only the loadings for test 2 differ significantly - and this only at the 5% level - in the replicated study. For factor II tests 4 and 8 differ at the same level.

The results bear out the conclusions reached in the last chapter where the 1950A and 1951A factors were compared by Ahmavaara's method.

Reverting to formula (8) it will be recalled that when it was employed a weighted sum of the 'specifics' for test (8) - the most reliable test - was used as a standard for scaling the factor loadings. The fact is that the choice of test is immaterial since the assumption made is that the ratios of the 'specifics' are assumed to be known. This was demonstrated by repeating the calculations choosing test 3 instead of test 8 and showing that the critical ratios obtained were identical.

Using Burt's Standard Error Formula.

For comparison with the results obtained above by Lawley's equations the standard errors of the factor loadings in the two studies calculated by Burt's empirical formula were used to find new critical ratios for the difference between each of the pairs of factor loadings. These are given in table XXIII. The standard errors themselves are not directly comparable as the figures given in table XXII and XXIIa are for the adjusted loadings, but the critical ratios are comparable. The latter show close agreement with the earlier results: the loadings for test 2 on factor I in the two studies again are significantly different this time at the 1% level, but the loading on factor II of test 4 and 8 which before were significantly different at the 5% level are now not significant. The above investigation tends to increase our confidence in Burt's S.E. formula and suggests that it may be preferable to the S.V. formula for standard errors of loadings.

Working on this conclusion attention ~~was~~ momentarily reverted to the results of the analyses reported in chapter III and a few cases of how loadings obtained in different studies compare were considered. Details of the results would be cumbersome and so are not given here but in general it was found that centroid loadings agreed very well with their maximum likelihood counterparts.

In table XXIV the maximum likelihood loadings and standard errors of factors I and II in the 1950A and 1950B studies are given side by side to facilitate a comparison of the change incurred in the former loadings as a result of adding variables 7 and 9 to the battery.

An examination of the data in table XXIV shows that the inclusion of variables 7 and 9, which are cognitive variables, tends to increase the amount of variance on the first factor extracted by the other cognitive variables, viz. 6, 8 and 10, at the expence of the amount extracted by the neurotic variables (1 - 5). The reverse is the case with factor II: for example the loading for variable 8 is no longer significantly greater than its standard error. The amount of variance extracted by the variables, as shown by the h^2 column of table X and table XII, however, has everywhere increased to a small extent - as one would expect - from the 1950A to the 1950B analysis.

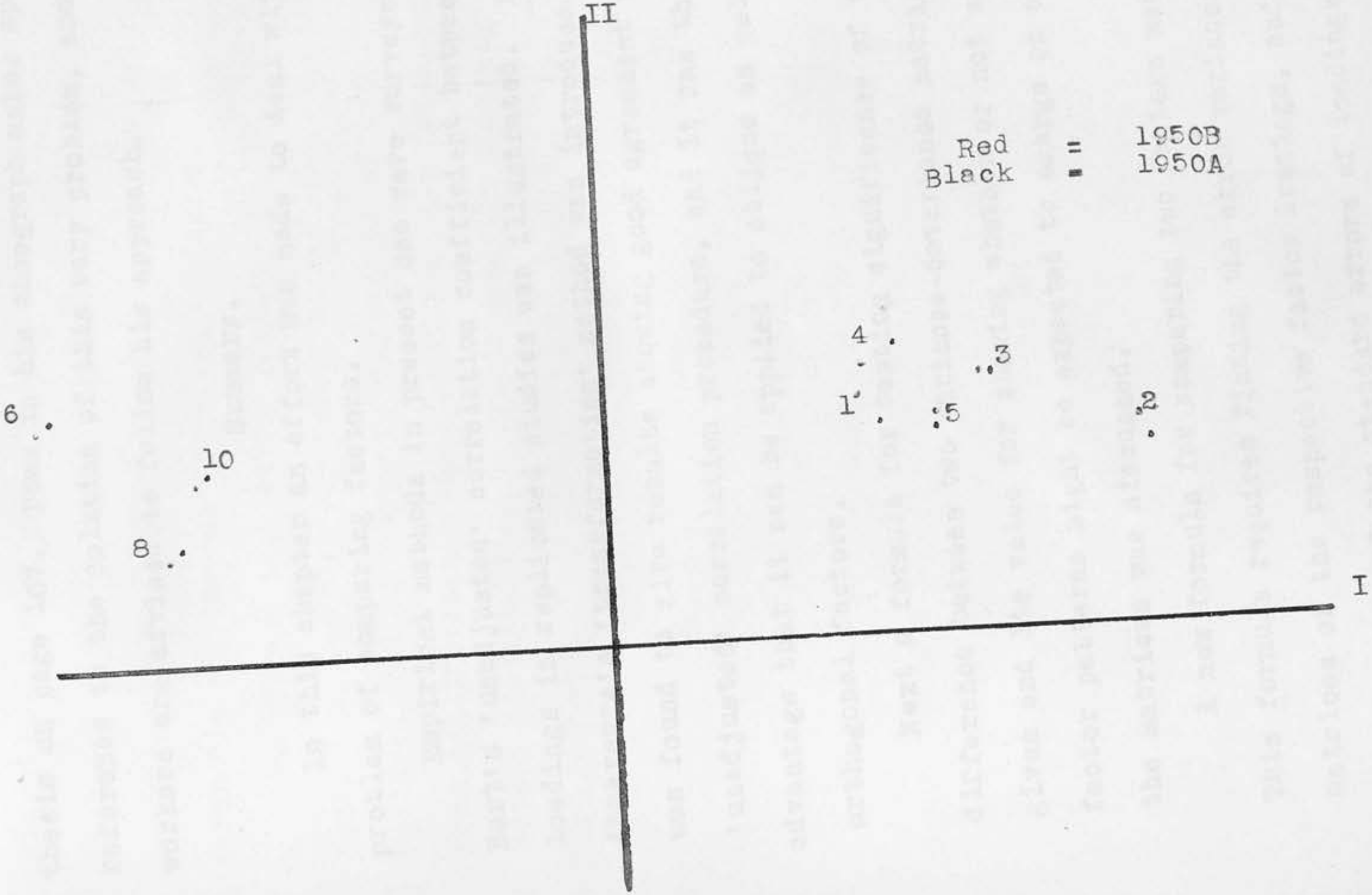
Figure 2 shows the plots of the loadings for the two analyses. Here it is seen that the effect of adding variables 7 and 9 is largely of a rotational nature. The relative positions of the variables have changed little. The latter type of invariance is what factor analysts stress, and their effort to demonstrate it from one study to the next is one justification for rotation procedures.

Discussion.

The writer feels that the above review of standard error formulae and their possible uses for comparing results obtained from replicated factor-analytic experiments hardly calls for celebration. The procedures which have a valid statistical basis require an exorbitant amount of computation, whilst possible empirical and less onerous alternatives do not inspire us with confidence. The chapter, however, need not end on a pessimistic note for it would appear that the recent work by Howe (1955) and Lawley (1956), referred to at the end of the last chapter, is directly relevant to the question of factorial invariance. Comparison of factor loadings obtained in one study with those obtained in a succeeding study is equivalent to comparison of the results from one sub-example with those from the other in the experimental design outlined in this

FIGURE II

Plots of Factors for the 1950A and 1950B Samples



thesis on page 107. Howe in his monograph makes special reference to the solution of this very problem, and future workers are advised to follow his approach.

Summary.

In this chapter an effort was made to deal with the problem of comparing factors.

Empirical methods in present use were surveyed. Firstly Burt's 'unadjusted' correlation coefficient between factor loadings in replicated studies was illustrated. Then Ahmavaara's 'transformation' method was introduced. It was found to give results showing good agreement with the 'unadjusted' correlation procedure, and it has the advantage that it can be applied to oblique as well as to orthogonal factors.

Next a formula for testing significance of the difference between two variance-covariance matrices was given and its value for showing whether or not similar factor patterns might be expected to emerge on analysis of the matrices was discussed.

A new formula for comparing two factors was given. This formula requires finding the error variance-covariance matrices of the respective factor loadings, and this led to a discussion of the standard errors of loadings. Formulae by Lawley and Burt were illustrated and the results they gave were found to show a resemblance. Finally, attention was given to the comparison of loadings from matrices which did not include identical variables.

An overall look at the findings in the chapter led to pessimism, but this was turned to optimism when it was realised that an answer to the problem of comparing results from replicated experiments had now been indicated in the recent contributions to factor analytic theory by Howe and Lawley.

Chapter VII

SUMMARY AND CONCLUSIONS

In this chapter an effort is made

1) to summarise the topics discussed in earlier chapters

2) to assess the numerous factorial techniques considered, and make what recommendations appear justified and practical in view of the evidence presented,

3) to try to clarify - if clarification is necessary - the role which factor analysis can play in psychological investigations.

Topics.

It was not the aim of this thesis to give a full and coherent account of factor analytic theory as it exists today. On the contrary the aim was to deal with points of controversy about the subject - to recount its imperfections and try to remedy its faults.

To get things into perspective a brief historical review of the growth of factor analysis was undertaken. This revealed one major source of confusion, to wit, the failure to distinguish clearly between the principal component model and the <u>factor</u> model proper. The relative claims of these two models on the psychologist's attention were then considered in some detail, when it was concluded that the principal component model was, in general, incompatible with his needs.

The plausibility of the factor concept which in essence claims that a p-dimensional complex of test variables can, in general, be accounted for in terms of m (less than p) dimensions, was then considered. Support for the idea was presented from the work of Spearman, Thomson, Thurstone and others.

At this point the customary word of caution about the danger of reifying factors was given.

Having completed a general introduction to factor analysis more technical topics were considered. Statistical methods for dealing with the factor model were discussed; the perennial question of 'how many factors' was surveyed and new recommendations made.

Following this a chapter was devoted to the interpretation of factors. Here the opportunity of discussing psychological as well as statistical aspects of the problem was taken. The fact that the statistician can answer the question only of the dimensionality of the factor space but cannot decide where to place the axes in it was emphasised. The job has to be completed by the psychologist, and as the aims of different studies differ, and the opinions of psychologists vary, it was admitted that one cannot hope to objectify completely interpretation procedures. However, a logical approach to adopt was outlined.

Finally, the question of factorial invariance was considered. A number of empirical methods for comparing factors derived from replicated experiments were described and illustrated. Though valuable, they were considered for statistical reasons not to be wholly satisfactory. A more rigorous approach to the comparison of factors using standard errors was then outlined and a new test for dealing with the problem proposed. Though this test was shown to have merit its use - on account of the lengthy computations involved - was not recommended, and the day was saved only by the realisation that recent work by Howe, and by Lawley, made possible a new and more satisfactory solution to the problem of comparing factors.

Factorial Techniques and Recommendations Concerning Their Use.

Having shown that the principal component model differed in its aims from the factor model and was of limited interest to the factor analyst, attention was focused on the factor problem proper. This is stated in precise statistical terms at the end of chapter II, where it is indicated that a valid statistical solution is provided by Lawley's maximum likelihood method.

It was considered beyond the scope of this thesis to give a detailed theoretical account of Lawley's solution, but some of its advantages are outlined at the end of chapter II. Chief of these are a) the fact that it provides a valid large sample test of the significance of residuals, and b) its unique quality amongst factorial techniques of giving results which are invariant under change of scale in the variables. (The latter advantage, though emphasised by Thomson, has not been sufficiently realised by other writers). In effect it means that loss of information which otherwise would result from the use of standardized scores can, with the maximum likelihood method, be obviated. In view of the dangers inherent in the standardization process (Uppsala Symposium) this advantage cannot be overstressed. In chapter III many examples of the use of the method - with details of the convergence of the estimates of the loadings with successive iterations - are given.

Lawley's solution, however, has not been accepted without criticism and reference to queries about it, by Kendall in particular, is made. These were found to be not of a serious nature - a fact which is supported in Howe's recent appraisal of Lawley's work. Moreover, since chapter III was written, a joint statement by Kendall and Lawley has done much to clarify the points at issue.

Some discussion of Lawley's approximate chi-square test followed. Theoretical evidence due to M.S. Bartlett, and other evidence from sampling investigation and practical examples shows that it is probably more accurate than has generally been claimed, at least for samples of 200 or more.

The disadvantage of the maximum likelihood method, in the absence of electronic computing facilities, is the amount of computation it involves, and convergence - in the iterative process - is known to be fairly slow. In view of this considerable attention was given to the centroid, or simple summation, method of analysis. Many practical examples of its use are also given in chapter III and the results compared with those obtained from the maximum likelihood analyses. Agreement is very good and - with reservations - it is safe to conclude that a centroid analysis in general gives a good approximation to the 'true' factor loadings. This conclusion has now been supported in an investigation by Lawley of the statistical <u>efficiency</u> of the centroid method.

One reservation about the method concerns the arbitrariness which attends the process of sign reflection. The limitations of the rules of thumb given by Thurstone are illustrated in the practical examples in chapter III. These rules do not necessarily ensure that maximum variance is extracted with succeeding factors and this limitation might well prejudice a test of the significance of residuals. Procedures by which the difficulty may, in general, be overcome are suggested in the thesis where the use of Holley's criterion is employed in conjunction with Thurstone's procedures.

Another approach to the sign reflection problem is that described by Thomson. It is satisfactory for small matrices. Burt too, and Guilford describe sign reflection procedures but these can readily get out of hand if the matrix being analysed is large.

Though the efficiency of the centroid method has now been established (Lawley, 1955) the problem of testing residuals obtained when using it has not yet been wholly resolved. In his article Lawley supplies a valid test for the particular case in which the specific variances of the tests employed are known. In practice, however, this information would seldom be available; but with efficient estimates of the factor loadings Lawley's 1940 approximate chi-square test would now be admissible. It would tend to overestimate the value of chi-square, but forewarned of this tendency the experimenter could use it with considerable confidence. He would, of course, be obliged to assure himself that he had obtained good convergence of his centroid loadings and this would call for one iteration at least.

The knowledge that residuals resulting from a centroid analysis can now be tested with some measure of statistical precision is satisfying, for, as was shown in chapter IV, the empirical tests customarily used for this purpose do not inspire confidence.

Having dealt with the problems of factor extraction and tests of significance the rotation techniques involved in factor interpretation were reviewed and evaluated. This forms the subject matter of chapter V of the thesis.

First the aims of rotation are discussed and a novel approach to the problem by Eysenck, called "criterion analysis", is described. Though the idea behind it is ingenious the method itself is of somewhat limited application and its <u>efficiency</u> as a statistical technique, as yet, an unknown quantity.

Leaving criterion analysis, the classical approach to the rotation problem, using <u>a priori</u> knowledge of the factor content of the tests, is illustrated by an example of Emmett's. The latter example is used to show the economy of description which a factor analysis of a matrix of correlations affords. In the first place a 9x9 matrix of correlations is replaced by a 9x3 matrix of loadings shown to be adequate to account for it. The latter matrix is then simplified further by rotation so that negative loadings are eliminated and many of the loadings themselves are reduced to zero. The result yields a concise and parsimonious description of the original matrix throwing the independent sources of covariation between the tests into clear outline.

But a priori knowledge of the factor composition of tests generally involves personal judgements and the rotation procedures to which it leads are, ipso facto, not wholly objective. There is nothing intolerable about this fact, nevertheless many factors theorists feel that it would be more satisfactory if rotation procedures could be objectified so that all experimenters, given the same set of data, would reach the same conclusions. To this end many analytical techniques, purporting to objectify the rotation procedure have been suggested. Two of these techniques, Thurstone's analytical method for 'simple structure', and Neuhaus and Wrigley's 'quadrimax method' were described and illustrated. In the nature of things the results they gave could not have been expected to be wholly satisfactory. This was so, but the methods were not altogether condemned as it was felt that they might well be helpful when trying to locate the main dimensions in exploratory and pilot studies. The fact that these analytical techniques completely ignore sampling issues, however, was deplored; negligence of this kind has in the past brought great discredit on factor analysis.

In chapter V, too, Thurstone's 'simple structure' concept was examined in detail. An understanding of this concept is essential, not just before the rotation of axes is undertaken, but (and this is what is so often forgotten) before the factorial experiment itself is designed. It must be emphasised without delay that difficulty was experienced with Thurstone's concept. From the statistical viewpoint it was found to be inadequately defined in the literature, and an honest effort by Tucker (1955) to meet the need resulted in a model which was so beset about by restrictions as to render it quite impractical. To replace it the writer recommends 'simple structures' (in the plural). No originality is claimed for the idea. By 'a simple structure' is meant a pattern of factor loadings in which every test does not necessarily contribute to every factor. The onus of defining the special simple structure he has in mind in any particular study is laid on the experimenter. He is obliged, having chosen his battery, to state the number and location of the zero loadings he expects on each of his rotated factors. If, from his antecedent knowledge of his tests, the experimenter cannot do this then it is suggested that he devide his sample randomly in half. One sub-sample can then be used for exploratory purposes to discover the most probable factor pattern, and this pattern can then be checked on the other sub-sample. The chief reason for advocating this procedure is that recent work by Howe (1955) and Lawley (1956) has shown how the check implied can be validly made.

Finally, in chapter VI, attention was given to possible techniques for dealing with the problem of factorial invariance and the comparison of results from different factorial experiments. Burt's 'uncorrected' correlation co-efficient method and Ahmavaara's transformation matrix method were critically examined and illustrated. Each was found to have merit, but a final decision about them had to be withheld owing to their empirical nature.

An attempt was then made to deal with a fairly straightforward problem of comparing factors from replicated experiments by orthodox statistical methods, making use of formulae for the standard errors of factor loadings. Some progress was made and details of the results are given on pages 129 and 130 and in tables XXI to XXIV. This approach at the moment, however, is not wholly satisfactory for the standard error formulae concerned tend to be very cumbersome and involve an exorbitant amount of calculation. On reflection it appears to the writer that an alternative, and more promising approach, lies in an adaptation of Howe's and of Lawley's more recent work. Following them, prior knowledge of the factor composition of tests in a battery would be used to postulate the pattern of loadings to be expected on completing an analysis. Factor invariance would then be bourne out to the extent that the pattern postulated was verified.

The Role of Factor Analysis in Psychology.

Although the main emphasis in this thesis is on the statistical aspects of factor analysis the writer was always conscious of an obligation to keep psychological considerations in the forefront of the discussion. On the psychological side it is no longer sufficient to give the customary warning to beginners about the dangers of reifying factors, or to discuss and try to clarify the logic behind 'causal explanations' (chapter II). The interpretation of factors involves practical issues, and throws up problems for the experimental psychologist of more immediate concern than metaphysical considerations. An effort to state clearly just what the writer has in mind is made in chapter V, which deals with the interpretation of factors. There some of the psychological ideas - as opposed to the psychometric speculations - which Thurstone, here and there, puts forth are recalled. He makes it quite clear that the isolation of so-called 'primary factors' is only the first step in the investigation he has in mind. Such factors must be further scrutinised, and if possible, investigated at the experimental level. More definite information, possibly of a neuro-physiological nature, must be sought to account for the particular structuring of human behaviour which the primary factors suggest. It was as a guide to such experimental work that - in Thurstone's view - the ultimate value of factor analysis lay. This point has since been reiterated by Eysenck.

Here it is well to look deeper and see what kind of evidence might be acceptable as support for a 'factor'. In this vein findings regarding localization of brain function were referred to (page 25). For example, Rosvold et al. from their experiments with monkeys write, "it is probable that one or more of the discrete structures in the temporal lobe are critical in bringing about alterations in aggressiveness".

If knowledge of association of this kind between structure and function in the brain becomes established, and if individual differences can be demonstrated, then it would be reasonable to expect evidence of such associations to be reflected in test performance, and in a subsequent factor analysis of the data.

Emperimental evidence of a different kind, which is considered to support the factor concept, is given by Drever when discussing the results of an experiment on early learning and space perception. "Certain perceptual abilities", he writes, "having to do with objects in space seem to require a long apprenticeship either in the visual or in the tactile-kinesthetic modalities, and .... once this apprenticeship has been served different amounts of later practice have no appreciable effect. We have in fact something rather like the kind of abilities identified by factorial studies of test performance".

On this note our discussion of factor analysis may conveniently end. Factors, it is concluded, are not an end in themselves but merely a means to an end. The procedure by which they are arrived at is so elaborate that it may, one day, be abandoned by the psychologist as uneconomical. Whether or no, the statistician will continue to be intrigued by the remarkable fact that a p-dimensional complex of test variates can, in general, be adequately represented by a space of m, less than p, dimensions.

TABLE III

Mr James's Correlation Matrix

Tests	1	2	3	4	5	6	7	8	9	10	11	12
1		345	594	404	579	-280	-449	033	-188	-303	-200	075
2	345		477	338	230	-159	-205	039	-120	-168	-145	141
3	594	477		498	505	-251	-377	033	-186	-273	-154	102
4	404	338	498		389	-168	-249	017	-173	-195	-055	009
5	579	230	505	389		-151	-285	-003	-129	-159	-079	037
6	-280	-159	-251	-168	-151		363	087	359	227	260	019
7	-449	-205	-377	-249	-285	363		-103	448	439	511	065
8	033	039	033	017	-003	087	-103		099	-226	-065	159
9	-188	-120	-186	-173	-129	359	448	099		429	316	093
10	-303	-168	-273	-195	-159	227	439	-226	429		301	007
11	-200	-145	-154	-055	-079	260	511	-065	316	301		-075
12	075	141	102	009	037	019	065	159	093	007	-075	
otal	610	773	968	815	934	306	158	070	948	079		632

TABLE IV

Mr James's Matrix - Centroid Loadings

	(a) I	(a) II	(b) (1) I	(b) (2) II	(b) (3) III	(b) (4) IV	(b) (5) V	(6) h^2	(7) Largest r	(8) Column Sum
1	343	-657	703	284	-168	208	-244	706	594	3.450
2	356	-431	494	258	158	-267	086	414	477	2.367
3	445	-643	703	381	-086	-061	078	657	594	3.450
4	374	-455	520	304	-127	-048	140	401	498	2.495
5	431	-478	542	358	-250	232	-070	543	579	2.546
6	191	472	-430	186	223	189	083	312	363	2.324
7	191	722	-673	368	155	-124	123	643	511	3.494
8	084	-113	124	049	422	330	076	311	226	0.864
9	398	542	-453	379	220	118	-159	436	448	2.540
10	148	579	-548	353	-115	-193	-272	549	439	2.727
11	321	533	-464	387	-109	067	257	447	511	2.161
12	225	-068	100	126	340	-073	-168	175	159	0.782
% var.	9.85	26.2	26.5	9.3	4.8	3.3	2.7	46.6		

TABLE V

(a)

Mr James's Matrix - Maximum Likelihood Loadings

Tests	Factor I		Factor II		Factor III	
	Iterations		Iterations		Iterations	
	(3)	(4)	(3)	(4)	(3)	(4)
1	758	758	222	222	-032	-022
2	460	460	209	209	137	115
3	733	735	323	324	036	029
4	526	526	259	259	-046	-053
5	583	583	325	325	-120	-106
6	-426	-426	219	221	177	185
7	-692	-691	394	394	-017	-026
8	072	072	-045	-045	502	527
9	-445	-444	442	442	201	208
10	-499	-497	318	316	-170	-164
11	-409	-408	441	440	-128	-126
12	051	051	148	148	328	315
% var.	26.6	26.6	9.1	9.1	4.3	4.4

(b)

Loadings for Five Maximum Likelihood Factors

	I	II	III	IV	V	h^2
1	762	214	-009	170	-088	663
2	469	193	096	-371	-066	408
3	745	297	019	-166	044	673
4	528	232	-046	-107	177	378
5	605	328	-088	251	000	544
6	-421	242	206	110	066	295
7	-680	418	-034	-088	088	654
8	071	-055	612	097	111	404
9	-430	465	224	078	-155	481
10	-487	339	-183	-002	-265	456
11	-394	451	-097	045	265	440
12	055	142	281	-146	-177	155
%var.	26.9	9.4	5.1	2.7	2.2	46.3

Table VI

Mr James's Matrix - Principal Component Loadings

Test	I	II	III	IV	h^2
			Components		
1	755	281	-048	-037	653
2	524	327	113	254	459
3	747	384	-030	016	707
4	584	377	-139	-175	533
5	605	415	-158	-142	583
6	-501	309	237	-323	507
7	-722	381	-043	068	673
8	087	000	792	-443	831
9	-518	537	205	-014	598
10	-570	374	-258	308	626
11	-468	506	-224	-279	603
12	065	258	607	608	809
%var.	30.9	13.6	10.6	8.0	63.2

Table VII

(a)

Mr Emmett's Matrix

Tests	1	2	3	4	5	6	7	8	9
1		525	395	471	346	426	576	434	639
2	523		479	506	418	462	547	283	644
3	395	479		355	270	254	452	218	504
4	471	506	355		691	791	443	285	505
5	346	418	270	691		679	382	149	409
6	426	426	254	791	679		372	314	472
7	576	547	452	443	382	372		385	680
8	434	283	218	285	149	314	385		470
9	639	644	504	505	409	472	680	470	
Total	3.810	3.862	2.927	4.047	3.344	3.770	3.837	2.538	4.323

(b)

Maximum Likelihood Factor Loadings

	Before Rotation				After Rotation		
Tests	I	II	III	h^2	I	II	III
1	668	313	111	557	639	071	379
2	693	241	-172	568	731	145	110
3	501	296	-236	394	628	000	000
4	840	-314	-028	805	532	698	186
5	704	-322	-143	619	463	635	033
6	804	-372	107	796	426	729	288
7	669	390	-058	602	739	005	237
8	448	266	409	439	329	-001	575
9	770	421	019	771	806	030	347
%var.	47.4	10.9	3.4	61.7	36.9	16.1	8.7

Principal Component Loadings

Tests	I	II	III	IV	h^2
1	748	-266	113	-256	709
2	761	-120	-263	-160	688
3	595	-293	-500	536	977
4	792	453	048	083	842
5	679	565	-059	001	784
6	748	509	177	071	855
7	755	-304	-081	-247	730
8	521	-366	681	303	961
9	833	-284	-016	-111	788
%var.	52.0	14.0	9.3	6.2	81.5

Table VIII

The Significance of the Latent Roots in a Principal Component Analysis

Root	d.f.	Chi-Square	Significance Level
L_1	28	$-201.5(-3.3830) = 681.67$	$P < 0.001$
L_2	27	$-201.5(-0.6331) = 127.57$	$P < 0.001$
L_3	20	$-201.5(-0.3245) = 65.39$	$P < 0.001$
L_4	14	$-201.5(-0.1404) = 28.29$	$P < 0.02$

Table IX
(a)
Mr Nigniewitsky's Correlation Matrix

Tests	2	3	4	5	6	7	8	9
1	638	704	017	139	114	150	160	228
2		570	193	097	238	245	196	269
3			015	179	102	081	007	172
4				616	543	-181	302	041
5					476	-372	100	-052
6						012	213	145
7							363	533
8								447

(b-
Centroid Loadings

	First Estimates				First Iteration			
Tests	I	II	III	h^2	I	II	III	h^2
1	634	434	337	704	628	427	342	694
2	673	310	213	594	669	305	218	588
3	552	418	447	679	552	418	459	690
4	480	-670	108	691	492	-686	091	721
5	389	-587	350	618	396	-595	349	633
6	514	-422	049	445	503	-407	046	421
7	272	391	-500	477	288	406	-522	520
8	505	-091	-475	489	501	-088	-464	474
9	509	214	-494	549	513	219	-497	558
%var.	26.6	18.2	13.5	58.3	26.6	18.5	13.8	58.9

Table X

CORRELATION MATRIX 1950 - STUDY (A)

Tests	1	2	3	4	5	6	8	10
1		475	297	134	205	-133	-222	-122
2	475		479	424	448	-517	-379	-399
3	297	479		453	381	-305	-334	-162
4	134	424	453		321	-151	-106	-190
5	205	448	381	321		-271	-294	-091
6	-133	-517	-305	-151	-271		570	568
8	-222	-379	-334	-106	-294	570		335
10	-122	-399	-162	-190	-091	568	335	

CENTROID LOADINGS

Tests	1st. estimates			1st iteration			2nd. iteration		
	I	II	h^2	I	II	h^2	I	II	h^2
1	398	199	198	388	186	185	387	182	183
2	814	141	682	825	157	705	834	164	722
3	650	306	516	635	289	487	632	280	478
4	466	325	323	456	318	309	455	314	306
5	521	232	325	507	217	304	505	210	299
6	-706	530	779	-715	548	812	-726	569	851
8	-590	290	432	-580	257	402	-576	242	390
10	-497	375	386	-489	356	366	-487	347	358
% Var.	35.3	10.2	45.5	34.8	9.8	44.6	35.0	9.8	44.8

MAXIMUM LIKELIHOOD LOADINGS

Tests	1st. factor iterations					2nd. factor iterations					h^2
	A	1	2	3	4	B	1	2	3	4	
1	430	395	385	376	370	159	242	269	291	307	231
2	800	792	783	777	773	142	171	215	244	265	668
3	610	591	572	559	549	365	357	365	371	377	444
4	460	437	423	410	401	373	398	406	413	418	336
5	525	498	486	477	471	246	269	282	294	303	314
6	-730	-739	-762	-778	-791	400	413	395	383	374	766
8	-580	-596	-602	-607	-610	246	219	199	182	169	401
10	-500	-541	-556	-564	-569	370	338	318	298	281	403
Var.		34.6	34.4	34.3	34.3		9.7	9.9	10.1	10.3	44.5

Table XI
CORRELATION MATRIX 1951 STUDY (A)

Tests		2	3	4	5	6	8	10
1		435	334	390	332	-130	-008	000
2	435		373	483	400	-270	-226	-148
3	334	373		519	218	-225	-067	-126
4	390	483	519		276	-095	024	-049
5	332	400	218	276		-214	-159	-025
6	-130	-270	-225	-095	-214		598	550
8	-008	-226	-067	024	-159	598		558
10	000	-148	-126	-049	-025	550	558	

CENTROID LOADINGS
following Thurstone's rules for sign reflection.

Tests	1st. estimates			1st. iteration		
	I	II	h^2	I	II	h^2
1	526	347	397	520	353	395
2	450	560	516	464	567	537
3	454	429	390	421	436	367
4	608	387	519	614	396	534
5	361	377	272	327	384	254
6	239	-720	576	235	-715	566
8	387	-674	604	393	-671	605
10	387	-629	545	388	-625	541
% Variance	19.3	28.4	47.7	18.8	28.7	47.5

CENTROID LOADINGS
using Holley's criterion for sign reflection and the communalities obtained by Thurstone's method.

Tests	I	II	h^2
1	475	390	378
2	675	273	530
3	524	263	344
4	546	472	521
5	441	193	232
6	-622	449	588
8	-516	585	608
10	-469	562	536
% Variance	29.0	17.7	46.7

Table XI continued

MAXIMUM LIKELIHOOD LOADINGS - 1951 STUDY

(using the Thurstone centroids as first guesses.)

	1st. factor iterations					2nd. factor iterations				
Tests	A	1	2	3	4	B	1	2	3	4
1	520	527	538	558	581	353	298	262	210	128
2	464	492	521	562	618	567	528	493	440	347
3	421	464	491	526	567	436	386	355	307	229
4	614	640	662	689	716	396	334	289	223	117
5	327	337	353	380	417	384	347	325	292	234
6	235	205	158	086	-026	-715	-738	-757	-772	-818
8	393	361	318	250	140	-671	-696	-722	-751	-770
10	388	334	285	218	116	-625	-642	-656	-675	-689
% Var.	18.8	19.3	19.6	20.5	22.0	28.7	27.4	26.9	24.9	26.0

The above iterations continued with new guesses.

Tests	C	5	6	D	5
1	600	592	557	040	-021
2	660	683	701	240	167
3	590	607	606	160	083
4	740	733	677	040	-074
5	440	457	474	140	122
6	000	-227	-461	-880	-829
8	070	-067	-287	-785	-756
10	060	-071	-268	-700	-678

Iterations repeated commencing with the Holley loadings.

	1st. factor iterations						2nd factor iterations						
Tests	E	1	2		6	7	F	1	2		6	7	h^2
1	475	428	394	-	346	340	390	425	457	-	482	486	352
2	675	640	611	-	569	563	273	333	366	-	426	432	504
3	524	494	471	-	432	427	263	340	393	-	433	439	375
4	546	491	451	-	394	387	472	525	556	-	612	618	532
5	441	425	407	-	381	377	193	230	264	-	288	292	227
6	-622	-662	-692	-	-722	-727	449	401	201	-	280	272	603
8	-516	-574	-613	-	-666	-671	585	535	496	-	441	433	638
10	-469	-526	-556	-	-592	-596	562	492	443	-	378	370	492
Var.	29.0	28.8	28.5		28.1	28.0	17.7	17.8	17.0		18.4	18.5	46.5

Table XII

CORRELATION MATRIX 1950 - STUDY (B)

Tests	1	2	3	4	5	6	7	8	9	10
1		475	297	134	205	-133	-137	-222	094	-122
2			479	424	448	-517	-095	-379	-075	-399
3				453	381	-305	-096	-334	-134	-162
4					321	-151	007	-106	-046	-190
5						-271	-238	-294	-072	-091
6							142	570	164	568
7								290	-066	-004
8									218	335
9										210

MAXIMUM LIKELIHOOD LOADINGS

	1st. Factor Iterations			2nd. Factor Iterations			Communality
Tests	A	1	2	B	1	2	h^2
1	370	361	359	307	354	350	251
2	773	765	761	265	280	299	669
3	549	543	536	377	363	376	429
4	401	389	382	418	400	404	309
5	471	470	467	303	332	324	323
6	-791	-801	-807	374	357	353	776
7	-200	-189	-187	-300	-046	-038	036
8	-610	-620	-622	169	137	145	408
9	-200	-189	-189	300	157	133	053
10	-569	-572	-576	281	297	274	407
% Var.	28.2	28.1	28.0	10.0	8.7	8.6	36.6

Table XIII

Correlation Matrix - 1951B Sample

Tests	2	3	4	5	6	7	8	9	10
1	435	334	390	332	-130	-162	-008	023	000
2		373	483	400	-270	-136	-226	-154	-148
3			519	218	-225	017	-067	-118	-126
4				276	-093	-01o	024	-096	-049
5					-214	-216	-159	-096	-025
6						533	598	532	550
7							419	282	276
8								532	558
9									518

Maximum Likelihood Loadings

	1st Factor Iter.s				2nd Factor Iter.s				3rd Factor Iter.s			
	A	1	2	3	B	1	2	3	C	1	2	3
1	-135	-175	-201	-218	512	582	580	571	275	-032	-142	-177
2	-311	-382	-410	-423	522	601	585	569	408	106	000	030
3	-219	-269	-287	-300	461	512	512	525	357	243	227	235
4	-142	-190	-215	-233	515	647	652	671	468	285	213	209
5	-200	-261	-287	-302	465	446	414	400	107	-164	-255	-264
6	765	823	815	-815	-288	016	076	101	190	182	128	100
7	495	527	523	536	-215	-002	071	101	295	417	355	396
8	714	729	731	729	-061	173	215	238	137	055	028	025
9	639	650	650	648	-060	147	173	185	081	-092	-156	-180
10	781	720	688	678	191	245	228	235	-058	-269	-241	-250

	1st Factor Iterations		2nd Factor Iterations		Communality h^2
	4	5	4	5	
1	-254	-255	544	543	360
2	-478	-479	524	527	507
3	-322	-344	498	500	362
4	-298	-295	649	654	515
5	-327	-328	355	354	233
6	639	748	158	148	581
7	472	487	148	139	256
8	727	732	313	314	634
9	640	640	243	236	465
10	656	653	285	275	502
%var.		27.7		16.5	44.15

Table XIV

Correlation Matrix - 1950C Sample

Tests	2	3	4	5	6	8	10	11	12	13	14	15
1	475	297	134	205	-133	-222	-122	-097	376	-083	222	240
2		479	424	448	-517	-379	-399	-291	302	-209	125	249
3			453	381	-305	-334	-162	-284	300	-283	055	141
4				321	-151	-106	-190	-123	075	-252	173	152
5					-271	-294	-091	-127	049	-191	047	088
6						570	568	182	-172	130	034	-136
8							335	254	-352	185	-086	-249
10								233	-063	051	-087	-192
11									-176	208	-293	-234
12										-204	069	235
13											-123	-172
14												694

Centroid Loadings

Tests	I	II	III	h^2
1	435	066	167	221
2	782	-159	276	713
3	616	078	364	518
4	444	059	281	280
5	437	-101	276	277
6	-612	678	196	873
8	-588	208	080	395
10	-449	336	268	386
11	-419	-081	108	194
12	387	131	057	170
13	-343	-150	-091	148
14	378	524	-407	583
15	527	409	-445	626
%var.	25.8	8.9	6.8	41.4

Table XV

Correlation Matrix - 1951C Sample

Tests	2	3	4	5	6	8	10	11	12	13	14	15
1	435	334	390	332	-130	-008	000	041	-019	-050	127	060
2		373	483	400	-270	-226	-148	-028	003	-074	066	046
3			519	218	-225	-067	-126	053	-017	050	102	056
4				276	-095	024	-049	-012	077	027	016	041
5					-214	-159	-025	-002	037	088	032	006
6						598	550	204	-164	031	-305	-220
8							558	287	-280	-182	-206	-177
10								512	-164	-001	-224	-185
11									-086	079	-221	-292
12										079	116	089
13											-148	-082
14												647

Centroid Loadings

	Following Thurstone's Rules		Applying Holley's Criterion followed by Thurstone's Rules			
Tests	I	II	I	II	III	h^2
1	347	479	405	236	325	325
2	302	626	568	147	227	396
3	300	517	417	185	179	240
4	433	574	449	309	292	382
5	243	444	344	149	081	147
6	376	-574	-656	552	198	774
8	462	-520	-545	607	376	807
10	526	-460	-567	678	221	830
11	415	-217	-327	501	-027	359
12	-166	166	169	-202	-117	083
13	028	004	-012	034	-184	035
14	-487	377	532	-506	378	682
15	-467	339	467	-496	407	630
%var.	14.1	19.7	20.5	16.6	6.7	43.8

Table XVI

Rotated Loadings for Emmett's Matrix got by Three Different Rotation Techniques

Variable	(a) Emmett's Interpretation			(b) Quadrimax Solution			(c) Thurstone's Analytical Simple Structure		
	I	II	III	I	II	III	I	II	III
1 Non-verbal	639	071	379	690	260	111	486	043	319
2 " "									
2 " "	731	145	110	656	328	-172	566	109	037
3 " "	628	000	000	562	152	-236	554	-032	-042
4 Verbal	532	698	186	361	821	-028	130	672	033
5 "	463	635	033	261	729	-143	116	612	-101
6 "	426	729	288	295	835	107	009	711	136
7 Spatial	739	005	237	746	207	-058	622	-030	183
8 "	329	-001	575	503	135	409	221	-010	544
9 "	806	030	347	839	257	-019	655	-007	283

Table XVII

Rotated Loadings for the 1950C and 1951C Matrices found by the Quadrimax and by Thurstone's Analytical Method

(a)
1950C

Variable	Quadrimax I	Quadrimax II	Quadrimax III	Thurstone I	Thurstone II	Thurstone III
1	-264	363	147	000	396	058
2	-653	531	085	-234	773	-049
3	-363	611	125	053	649	-020
4	-257	456	081	050	479	-026
5	-353	391	-030	-079	516	-123
6	924	118	081	836	-489	080
8	592	-135	-169	393	-417	-117
10	597	129	-123	547	-238	-130
11	286	-138	-306	122	-200	-261
12	-203	275	233	012	273	162
13	153	-288	-206	-058	-258	-135
14	-023	037	764	095	-150	735
15	-212	050	762	-053	-025	725

(b)
1951 C

Variable	Quadrimax I	Quadrimax II	Quadrimax III	Thurstone I	Thurstone II	Thurstone III
1	-016	555	131	053	508	068
2	-212	590	063	-151	562	039
3	-098	480	023	-071	463	-011
4	-003	615	063	034	582	-007
5	-101	369	-039	-104	369	-058
6	874	-103	013	772	-126	-157
8	886	075	132	841	021	-064
10	909	038	-035	783	024	-227
11	562	023	-206	396	058	-319
12	-284	-031	-035	-267	-015	028
13	-016	-060	-177	-099	-017	-162
14	-610	261	493	-298	151	576
15	-552	229	522	-233	112	597

Table XVIII (a)

Results Got by Burt's Group Factor Method

using an iterative method for obtaining the group factor loadings.

Variable	1950A Factors I	II	III	h^2	1951A Factors I	II	III	h^2
1	284	318		182	160	672		477
2	771	438		786	747	367		693
3	477	487		465	485	340		351
4	266	511		332	139	841		727
5	390	377		294	462	236		269
6	-630		638	804	-469		683	686
8	-610		253	436	-219		735	588
10	-441		406	359	-175		707	530
var.%	26.1	11.7	7.9	45.7	16.9	18.3	18.8	54.0

(b)

Quadrimax Results

	I	II	III	h^2	I	II	III	h^2
1	113	467		231	594	-010		352
2	461	675		668	674	-225		504
3	215	631		444	603	-109		375
4	072	575		336	730	025		532
5	197	524		314	453	-149		227
6	-859	-174		766	-181	756		603
8	-590	-229		401	-016	799		638
10	-624	-115		403	-027	702		492
%var.	22.3	22.2		44.3	24.3	22.3		46.5

Table XIX

Standard Errors of Factor Loadings by Burt's Formula
and by the S.V. Formula

Maximum Likelihood Results

1950A Sample

Test	Factor II	Burt's St. Error	S.V. St. Error
1	.307	.1024	.0860
2	.265	.1053	.0565
3	.377	.0971	.0731
4	.418	.0934	.0800
5	.303	.1028	.0812
6	.374	.0974	.0474
8	.169	.1099	.0759
10	.281	.1043	.0758

1951A Sample

Test	Factor II	Burt's St. Error	S.V. St. Error
1	.486	.0794	.0662
2	.432	.0845	.0579
3	.439	.0839	.0650
4	.618	.0643	.0562
5	.292	.0951	.0723
6	.272	.0963	.0518
8	.433	.0845	.0495
10	.370	.0897	.0586

Table XX

Adjusted Factor Loadings
(the specific variance of each test taken as 0.460)

	1950A		1951A	
Tests	Factors		Factors	
	I	II	I	II
1	.286	.237	.289	.410
2	.910	.312	.542	.416
3	.499	.343	.366	.377
4	.334	.348	.384	.613
5	.386	.248	.291	.225
6	-1.109	.524	-.783	.293
8	-0.535	.148	-.756	.488
10	-0.500	.247	-.567	.352

Table XXI

Variances and Covariances of Factor Loadings

1950A Factor I

Tests 1	2	3	4	5	6	8	10
.0069	.0042	.0025	.0015	.0016	.0021	-.0002	.0008
	.0115	.0043	.0042	.0036	-.0008	-.0006	-.0006
		.0093	.0044	.0031	.0017	-.0017	.0013
			.0087	.0028	.0029	.0010	.0007
				.0074	.0011	-.0005	.0013
					.0179	.0051	.0060
						.0070	.0017
							.0078

1951A Factor I

1	2	3	4	5	6	8	10
.0153	.0122	.0106	.0169	.0068	.0073	.0128	.0093
	.0163	.0108	.0178	.0072	.0071	.0121	.0091
		.0129	.0165	.0056	.0063	.0117	.0081
			.0295	.0093	.0113	.0193	.0135
				.0074	.0036	.0065	.0054
					.0115	.0114	.0084
						.0215	.0129
							.0131

Table XXII

The Significance of the Differences between Factor Loadings in the 1950A and 1951A Analyses

Test	Factor I Diff.s between Loadings	S.E. of diff.	C.R.	Sig. Level
1	.000	.149	.000	N.S.
2	.368	.167	2.207	5%
3	.133	.149	0.893	N.S.
4	.050	.195	0.256	"
5	.095	.122	0.781	"
6	.326	.172	1.901	"
8	.221	.169	1.309	"
10	.067	.145	0.463	"

Table XXIIa

	Factor II Diff.s between Loadings	S.E. of diff.	C.R	Sig. Level
1	.173	.117	1.478	N.S.
2	.104	.138	0.753	"
3	.034	.119	0.285	"
4	.265	.129	2.057	5%
5	.023	.116	0.198	N.S.
6	.231	.163	1.414	"
8	.340	.159	2.133	5%
10	.105	.139	0.760	N.S.

Table XXIII

Differences between the 1950A and 1951A Factor Loadings using Burt's Standard Error Formula

Maximum Likelihood Loadings

Test	Factor I Diff. bet. Loadings	S.E. of diff.	Sig. of diff.	Factor II Diff. bet. Loadings	S.E. of diff.	Sig. of diff.
1	.030	.129	N.S.	.179	.130	N.S.
2	.210	.079	1%	.167	.135	"
3	.122	.111	N.S.	.062	.129	"
4	.041	.125	"	.200	.114	"
5	.094	.120	"	.011	.140	"
6	.064	.062	"	.102	.137	"
8	.061	.088	"	.264	.139	"
10	.027	.098	"	.089	.138	"

Centroid Loadings

Test	Factor I Diff. bet. Loadings	S.E. of diff.	Sig. of diff.	Factor II Diff. bet. Loadings	S.E. of diff.	Sig. of diff.
1	.088	.121	N.S.	.208	.141	N.S.
2	.159	.062	1%	.109	.146	"
3	.108	.097	N.S.	.017	.143	"
4	.091	.112	"	.158	.130	"
5	.064	.114	"	.017	.147	"
6	.104	.078	"	.120	.113	"
8	.060	.104	"	.343	.126	"
10	.018	.114	"	.215	.123	"

Table XXIV

The Maximum Likelihood Loadings and their Standard Errors

for the 1950A and 1950B Samples

Test	Factor I				Factor II			
	1950A		1950B		1950A		1950B	
	Loading	S.E.	Loading	S.E.	Loading	S.E.	Loading	S.E.
1	.370	.098	.359	.096	.307	.102	.350	.096
2	.773	.046	.765	.046	.265	.105	.299	.100
3	.549	.079	.536	.078	.377	.097	.376	.094
4	.401	.095	.382	.094	.418	.093	.404	.092
5	.471	.088	.467	.086	.303	.103	.324	.098
6	-.791	.042	-.807	.038	.374	.097	.353	.096
7			-.187	.106			-.038	.110
8	-.610	.071	-.620	.067	.169	.110	.145	.107
9			-.189	.106			.133	.108
10	-.569	.077	-.576	.073	.281	.104	.274	.101

APPENDIX A

Worries and Anxieties

We all have a good many things that we are anxious about and that worry us in some way. Read through the list of words given below as carefully as you can. As you read through them underline all the things that have ever worried you or that still worry you now. It does not matter how many you underline. What you must be sure of is that you underline everything that has ever worried you.

---

1. Crowds Lifts Wireless Borrowing Sickness
2. Death Whispering Neighbours Daydreaming Blame
3. Teacher Death Work Food Books
4. Failure Stammering Home Accidents Sleep
5. Burglars Stealing Tiredness Clubs School
6. Shouting Moving Dizziness Forgetfulness Rudeness

Word Likes And Dislikes

Here is a list of words. Some of these you may like, others not. We all differ with regard to the kind of words we like, and it is interesting to see <u>your</u> choice. Put <u>L</u> after each word you like and <u>D</u> after each word you dislike. The reasons for which you like or dislike a word do not matter, just remember to put

L after each word you like

D after each word you dislike.

Try to work as quickly as you can, thinking carefully about each word.

---

1. Food Ghosts Fireworks School Train
2. Girl Potatoes Mother Meat Nurse
3. Noise Doctor Man Stranger Dreams
4. Sister Milk Dark Animals Brother
5. Work Bed Boy Father Woman
6. Night Baby Dinner Teacher Lesson

Interests.

Read through the list of words given below and underline those that you like or that you are interested in. It does not matter how many you are interested in. However, be sure you only underline those things that interest you.

---

1. Lectures Riding Boxing Shopping Stamp-collecting
2. Netball Film-stars Swimming Rounders Debating
3. Argueing Horses Betting Gossip Running
4. Buses Pretty girls Sports Outings Films
5. Carpentry Brassbands Camping Fairs Acting
6. Teaching Sewing Music-halls Snooker Processions

Sentence Completion Test

Here are a list of sentences which are incomplete. You are asked to finish the sentences any way you like. Don't worry about spelling. Put down the first thing that comes into your head.

1. Jane thought that some of the girls she knew ......

2. When I dream ......

3. The children were talking about the teacher; John said ....

4. Henry looked at his brothers and sisters and thought ......

5. The children at school ......

6. I often feel .....

7. Father, mother and the children were sitting at the table; Father said ....

8. Mary was thinking about school; she thought that ......

9. Ann was talking to the teacher; the teacher said .....

10. When I am older I wonder if I shall ......

11. James said that the other boys ......

12. Betty said, " My mother is always ....

QUESTIONAIRRE

Here is a list of questions about yourself. Read each carefully and answer them by underlining the word

YES if the answer is yes

NO if the answer is no.

Answer every question, do not leave any of them out.

1.	Are you often away from school through illness?	YES	NO
2.	Do you wish you did not get tired so easily?	YES	NO
3.	Do you think you have a poor appetite?	YES	NO
4.	Would you like to have more friends?	YES	NO
5.	Do you find that you think a lot about what people say to you?	YES	NO
6.	When you go to bed do you find it difficult to go to sleep?	YES	NO
7.	Does your attention wander so badly that you lose the thread of what you are doing?	YES	NO
8.	Do you like to play alone better than with other children?	YES	NO
9.	Do you sometimes feel happy, sometimes sad, without knowing why?	YES	NO
10.	Do you often have bad dreams?	YES	NO
11.	Do you worry about things which may happen to you?	YES	NO
12.	Do you find it difficult getting used to something new (like moving to a new school)?	YES	NO
13.	Do you feel that others in your class are better than you are in most things?	YES	NO
14.	Do you jump or start at sudden or loud noises?	YES	NO

Questionnaire continued -

15.	Would you like to play more with other children than you do?	YES	NO
16.	Have you ever stammered and stuttered?	YES	NO
17.	Do your classmates often quarrel with you?	YES	NO
18.	Do you often get headaches?	YES	NO
19.	Do you feel you cannot stick at a thing once you have started it?	YES	NO
20.	If someone upsets you do you feel cross for a long time?	YES	NO
21.	Do you often have aches and pains?	YES	NO
22.	Do you often lose your temper?	YES	NO
23.	Do you like being alone?	YES	NO
24.	Do you ever feel dizzy?	YES	NO

ANNOYANCES

Here are a number of things which tend to annoy many of us. Some of them probably annoy you.

Put a tick against the number of each item which you think would annoy you, if it happened to you. You may tick as many or as few items as you wish. Just be sure to tick everything that annoys you.

1. Harry was going up stairs when he slipped.
2. Jane could hear the windows rattling all the time.
3. David found some dirt in his food.
4. Pauline scraped her knife noisily on her plate.
5. Jack had forgotten what he was looking for.
6. Marry couldn't find her rubber anywhere.
7. Ernest was just becoming interested in what he was doing when he was stopped.
8. Sally found that she had far too many things to do.
9. John was dressing in a hurry when his shoelace broke.
10. Catherine found that the others were not listening to what she was saying.
11. Tony was unable to go off to sleep.
12. Ruth couldn't remember the name which was just on the tip of her tongue.
13. James was having to write on a very wobbly table.
14. Margaret had to wait a long time for the bus.
15. Peter was travelling in a crowded railway carriage.
16. Molly upset the tea all down her clothes.

WAYS TO BE DIFFERENT

A group of pupils were talking together. Most of them wished to be different in some way.

Here is a list of the ways in which they wanted to be different. You will probably find that you have made quite a number of these wishes yourself. If you would like to be different in any of these ways, put a tick next to those ways you agree with.

1. John wanted to be better at doing things with his hands.
2. Sally wanted to be better looking.
3. Tony wanted to have more things of his own.
4. Margaret wanted other children to like her more.
5. Jack wanted to be younger.
6. Mary wanted to stand up better for herself.
7. James wanted his mother and father to love him more.
8. Catherine wanted to have more money to spend.
9. Harry wanted to do better at lessons.
10. Pauline wished she did not lose her temper so easily.
11. Peter wanted his family to be nicer to him.
12. Ruth wanted to have more friends.
13. David wanted to be stronger.
14. Jane wished she were not so afraid.
15. Ernest wanted to have a nicer house.
16. Joan wanted to get along better with her mother and father.
17. Albert wanted to be better at games.
18. Betty wished other boys and girls would not tease her so much.

'Ways to be different' continued -

BUT most of all I WANT ....

(this is one you are
to make up yourself)

Now that you are finished go back to each question and decide how much you would like to be different.

If you want to be different only a little - put 1 next to it.

If you want to be different quite a lot - put 2 next to it.

If you want to be different very much - put 3 next to it.

B I B L I O G R A P H Y

AHMAVAARA, Y. 1954. Transformation analysis of factorial data, Annales Academiae Scientiarum Fennicae, Helsinki.

ANDERSON, T.W., and RUBIN, H. 1955. Statistical inference in factor analysis, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Stanford Univ. California.

BANKS, C. and BURT, C. 1954. The reduced correlation matrix, Br. J. Stat. Psychol. VII, 107 - 117.

BANNATYNE, A.D. 1953. An experimental study of introversion extraversion by means of projective techniques, Ph.D. Thesis, London University, (unpublished).

BARLOW, J.A. and BURT, C. 1954. The identification of factors from different experiments, Br. J. Stat. Psychol. VII, 52-56.

BARTLETT, M.S. 1937-38. The statistical conception of mental factors, Br.J.Psychol.(Gen.Sect.), XXVIII, 28, 97 - 104.

BARTLETT, M.S. 1950. Tests of significance in factor analysis, Br. J. Psychol.(Stat.Sect,), III, 77 - 85.

BARTLETT, M.S. 1951. A further note on tests of significance in factor analysis, Br.J.Psychol. (Stat.Sect.), IV, 1 - 2.

BARTLETT, M.S. 1953. Factor analysis in psychology as a statistician sees it, Nordisk Psykologi's Monograph Series, No.3, 23-34.

BENNETT, E. 1945. Some tests for the discrimination of neurotic and normal subjects, Br.J.Med.Psychol. XX, 271-277.

BOX, G.E.P. 1949. A general distribution theory for a class of likelihood criteria, Biometrika, XXXVI, 317-334.

BURT, C. 1940. <u>The Factors of the MInd</u>, London, Univ.Press, 338-40.

BURT, C. 1949. Subdivided factors, Br. J. Psychol.(Stat.Sect.) II, 41-63.

BURT, C. 1949. Alternative methods of factor analysis and their relations to Pearson's method of 'Principal Axes', Br. J.Psychol (Stat.Sect.), II, (98-121).

BURT, C. 1950. Group factor analysis, Br.J.Psychol. (Stat.Sect.), III, p.40.

BURT, C. 1915, Ann. Rep. Brit. Assoc. pp. 694 f.

BURT. C 1947-48, Factorial Study of Temperamental Traits, B. J. Psychol. (Stat. Sec) I, 178 - 203.

BURT, C. 1952. Tests of significance in factor analysis, Br.J.Stat.Psychol., V, 109-133.

CARROLL, J.B. 1953. An analytical solution for approximating simple structure in factor analysis, Psychometrika, XVIII, 23-38.

CATTELL, R.B. 1947. Confirmation and clarification of primary personality factors, Psychometrika, XII, 197-220.

DELGADO, J.M.R. 1952. Responses evoked in waking cat by eliciting stimulation of the motor cortex, Am. J. Physiol., CLXXI, Am.J.Physiol.CLXXI, 436-46.

DREVER, J. 1955. Early learning and the perception of space, Am. J. Psychol., LXVIII, 605-614.

EMMETT, W.G. 1949. Evidence of a space factor at 11 plus and earlier, Br.J.Psychol.(Stat.Sect.), II, 3-16.

Eysenck,H.J. 1947. <u>Dimensions of Personality</u>, London, Kegan Paul, Ltd.

EYSENCK, H.J. 1950. Criterion Analysis, Psychol. Rev.,LVII, 38-53.

EYSENCK, H.J. 1952. <u>The Scientific Study of Personality</u>, London, Routledge and Kegan Paul Ltd.

EYSENCK, H.J. 1953. The logical basis of factor analysis, Amer. Psychologist, VIII, 105-114.

EYSENCK, H.J. 1955. A dynamic theory of anxiety and hysteria, J.Ment. Sc., CI, 28-51.

FERGUSON, G.A. 1954. The concept of parsimony in factor analysis, Psychpmetrika, XVIIII, 281-290.

FISHER,R. 1953. The expansion of statistics, J.Roy.Stat.Soc., (A),CXVI, 1-10.

GUILFORD, J.P. and LACEY, S. 1947. The significance of residuals in factor analysis, U.S.Air Force Report.

HENRYSSON, S. 1950. The significance of factor loadings, Br. J.Stat.Psychol., III, 159-165.

HIMMELWEIT, H.T. and PETRIE, A. 1951. The measurement of personality in children, Br.J.Ed.Psychol., XXI, 9-29.

HOGBEN, L. 1955. Choice and Chance, VOL.II, London, Hodder and Stoughton.

HOLLEY, J.W. 1947. A note on the reflection of signs in the extraction of centroid factors, Psychometrika, XII, 263-265.

HOTELLING, H. 1933. Analysis of a complex of statistical variables into principal components, J.Ed.Psychol. XXIV. 417-441, 498-520.

HOWE, W.G. 1955. Some contributions to factor analysis, Report No. ORNL-1919, Oak Ridge National Laboratory.

HULL, C.L. 1943, Principles of behavior, New York, D. Appleton-Century.

HUMPHREY, G. 1951. Thinking, London, Methuen and Co. Ltd.

KENDALL, M.G. and SMITH, B.B. Factor analysis, J.Roy.Stat. Soc. B 1950, XII, 60-94.

Kendall, M.G. 1951. Regression, structure and functional relationship, Biometrika, XXXVIII, 11-25.

KENDALL, M.G. 1954. Review of Uppsala Symposium on Psychological factor analysis, J.Roy.Stat.Soc.(A), CXVII, 462-483.

KELLEY, T.L. 1927. Statistical Method, New York, Macmillan.

KELLEY, T.L. 1927. Crossroads in the Mind of Man, New York, Macmillan.

LAWLEY, D.N. 1940. The estimation of factor loadings by the method of maximum likelihood, Proc.Roy.Soc.Eding., LX, 64-82.

LAWLEY, D.N. 1949. Problems in factor analysis, Proc.Roy.Soc. Edingh., LXII.

LAWLEY, D.N. 1953. A modified method of estimation in factor analysis and some large sample results, Uppsala Symposium, Nordisk Psykologi's Monograph Series No3.

LAWLEY; D.N. and SWANSON, Z. 1954. Tests of significance in a factor analysis of artificial data, Br.J.Stat.Psychol., VII, 75-79.

LAWLEY, D.N. 1955. A statistical examination of the centroid method, Proc.Roy.Soc.Edingh.(Sect.A),LXIV,Pt.2, 175-189.

LAWLEY,D.N. 1957. Estimation in factor analysis under various initial assumptions, Br.J.Stat.Psychol., ~~VIII~~, (in press).

MAXWELL, A.E. 1955. Review of 'Transformation analysis of factorial data', by Y.Ahmavaada, Br.J.Psychol. XLVII.

MAXWELL, A.E. 1955. A statistical review of 'Psychical symptoms in neurological disease', by E.Busch, ACTA PSYCH. ET NEUR. SCAND., XXX, 483-497.

MAXWELL, A.E. 1956. Factor models and the factor concept, Am.J.Ed.Psychol.,XVI, (in press).

MOSIER, C.I. 1939. Influence of chance error on simple structure, Psychometrika, IV, 33-44.

MOURSY, E.M. 1952. The hierarchical organisation of cognitive levels, Br. J. Stat.Psychol., V, 151-180.

PEARSON, K. 1901. On lines and planes of closest fit to a system of points in space, Phil. Mag. II, 6th series, 557-72.

ROSVOULD,H.E., MIRSKY,A.F. and PRIBRAM,K.H. 1954. Influence of amygdalectomy on social behaviour in monkeys, J.Comp.Physiol. Psychol. XLVII, 173-178.

SAUNDERS, R.D. 1953. An analytic method for rotation to simple structure, Bull. RB-53-10, Ed. Test.Serv., Princeton.

SPEARMAN, C. 1904, General Intelligence objectively determined, Am. J. Psychol., 15, 201-293

TAYLOR, C.W. 1947. A factorial study of fluency in writing, Psychometrika, XII, 239-262.

THOMSON, G.H. 1916. A hierarchy without a general factor, Br. J.Psychol. VIII, 271-81.

THOMSON, G.H. 1939. The Factorial Analysis of Human Ability, London, Univ. of Lond.Press, (5th ed. 1951).

THOMSON, G.H. 1939a. The nature of the mind's factors, Pres. Address to Psychol. Sect. of the Br. Ass.

14

THORPE, J.G. 1953. A sociometric study of London school children, Ph.D. Thesis, London, (unpublished).

THURSTONE, L.L. 1938. Primary mental abilities, Psychom. Monog., No.1, Chicago, Univ. of Chicago Press.

THURSTONE, LL. 1947. Multiple Factor Analysis, Chicago, Univ. of Chicago Press.

THURSTONE, L.L. 1949. A factorial study of perception, Chicago, Univ.of Chicago press.

THURSTONE, L.L. 1954. Analytical method for simple structure, Psychometrika, XIX, 173-182.

TROUTON, D.S. and MAXWELL, A.E. 1955. The relation between neurosis and psychosis, J.Ment.Sc., CII, 1-21.

TUCKER, L.R. 1955. The objective definition of simple structure, Psychometrika, XX, 209-225.

VERNON, P.E. 1947. How many factors?, (Privately circulated manuscript).

VERNON, P.E. 1950. The Structure of Human Abilities, London, Methuen and Co. Ltd.

VERNON, P.E. 1955. The psychology of intelligence and G, Bull. Br. Psychol. Soc., No26, 1-14.

WARBURTON, F.W. 1954. The full factor analysis, Br. J. Stat. Psychol., VII, 101-106.

WILKS, S.S. 1932. Certain generalizations in the analysis of variance, Biometrika, XXIV, 471-494.

WILSON, R.C., GUILFORD, J.P., CHRISTENSEN, P.R. and LEWIS, D.J. 1954. A factor -analytic study of creative-thinking abilities, Psychometrika, XIX, 297-311.

WITTENBORN, J.R. 1951. Symptom patterns in a group of mental hospital patients, J. Consul. Psychol., XV, 290-302.

WOODROW, H. 1939. The common factors in 52 mental tests, Psychometrika, IV, 99-108.

WRIGLEY, C. and NEUHAUS, J.O. 1952. A re-factorization of the Burt-Pearson matrix with the Ordvac Electronic Computer, Br. J. Stat. Psychol., V, 105-108.

WRIGLEY, C. and NEUHAUS, J.O. 1955. The use of an electronic computer in principal axes factor analysis, J.Ed.Psychol., XLVI, 31-41.

WRIGLEY, C. and NEUHAUS, J.O. 1954. The minimax method: an analytical approach to orthogonal simple structure, Br.J. Stat. psychol. VII.